

Orð og tunga

15

Ritstjóri

Ásta Svavarsdóttir
Stofnun Árna Magnússonar í íslenskum fræðum
Neshaga 16
107 Reykjavík
sími 525 4437
asta@hi.is

Ráðgefandi ritnefnd

Anna Helga Hannesdóttir, Háskólanum í Gautaborg
Ari Páll Kristinsson, Stofnun Árna Magnússonar í íslenskum fræðum
Camilla Wide, Háskólanum í Turku
Christopher Sanders, ONP, Háskólanum í Kaupmannahöfn
Erla Erlendsdóttir, Háskóla Íslands
Finnur Friðriksson, Háskólanum á Akureyri
Guðrún Kvaran, Stofnun Árna Magnússonar í íslenskum fræðum
Hrafn Loftsson, Háskólanum í Reykjavík
Kendra Willson, UCLA, Los Angeles
Kristján Árnason, Háskóla Íslands
Lars Vikør, Háskólanum í Osló
Liisa Theilgaard, Háskólanum í Kaupmannahöfn
Margrét Jónsdóttir, Háskóla Íslands
Silvie Cinková, Univerzita Karlova, Prag
Veturliði Óskarsson, Háskólanum í Uppsölum
Zakaris Svabo Hansen, Fróðskaparsetri Færeyja, Þórshöfn
Þorsteinn G. Indriðason, Háskólanum í Bergen

Orð og tunga birtir fræðigreinar um mál og málfræði með áherslu á greinar sem snerta orðaforðann, þ. á m. nöfn og íðorð, og greinar um orðabókafræði og orðabókagerð. Í hverju hefti eru annars vegar greinar um fyrirfram ákveðið þema og hins vegar aðrar greinar á efnissviði tímaritsins almennt. Auk þess eru birtir ritdómar, bókafrægnir og frásagnir af rannsóknarverkefnum og ráðstefnum. Greinar skulu að jafnaði ritaðar á íslensku með útdrætti á ensku en einnig eru birtar greinar á norðurlandamálum og ensku með íslenskum útdrætti. Allar greinar í tímaritinu eru ritryndar þannig að auk ritstjóra lesa a.m.k. tveir ónafngreindir sérfræðingar hverja grein. Frekari upplýsingar um ritstjórnarstefnu, skil og frágang greina má finna á vefsíðu tímaritsins: http://www.arnastofnun.is/page/timarit_ot.

Orð og tunga

15

Ritstjóri
Ásta Svavarsdóttir



Stofnun Árna Magnússonar í íslenskum fræðum
Reykjavík 2013

© Stofnun Árna Magnússonar í íslenskum fræðum, 2013.

Öll réttindi áskilin.

Bók þessa má eigi afrita með neinum hætti, svo sem ljósmyndun, prentun, hljóðritun eða á annan sambærilegan hátt, að hluta eða í heild, án skriflegs leyfis höfunda og útgefanda.

ISSN 1022-4610

Umbrot: Bessi Aðalsteinsson.
Hönnun kápu: Björg Vilhjálmsdóttir.
Prentun og bókband: Leturprent ehf.

Efnisyfirlit

Formáli ritstjóra	vii
Þemagreinar:	
Jón Hilmar Jónsson: Að byggja undir íslensk-erlenda orðabók. Forgreining og orðabókarefni	1
Kristín Bjarnadóttir: Hvert á að sækja orðaforðann í orðabók?	23
Þórdís Úlfarsdóttir: ISLEX – norræn margmála orðabók	41
Rósa Elín Davíðsdóttir: Hlutverk tvímála orðabóka. Ólíkar notendaparfir í íslensk-frönsku ljósi	73
Aðrar greinar:	
Guðrún Kvaran: Orðabókarstörf Konráðs Gíslasonar	95
Susan M. Arthur: Are Oranges Yellow? Appelsínugulur as a Basic Color Term in Icelandic	121
Kirsten Wolf: Basic Color Terms in Old Norse-Icelandic: A Quantitative Study	141
Ritdómur:	
Auður Lorenzo: Íslensk-spænsk orðabók	163
Bókafregnir	169
Ráðstefnur 2013	179



Formáli ritstjóra

Í þessu hefti *Orðs og tungu* er sjónum einkum beint að íslensku sem viðfangsmáli í tvímála orðabókum. Vorið 2011 var haldið málþing á vegum tímaritsins undir yfirskriftinni *Íslenska sem viðfangsmál í íslensk-erlendum orðabókum. Sjónarmið og aðferðir við öflun, val og framsetningu efnisins*. Fimm fyrirlesarar, sem boðið var að taka þátt í þinginu, ræddu þemað frá ýmsum hliðum. Þrír þeirra fjölluðu um efnið út frá reynslu af gerð nýlegra íslensk-erlendra orðabóka: Guðrún H. Tulinius (*Íslensk-spænsk orðabók*), Stanisław J. Bartoszek (*Íslensk-pólsk, pólsk-íslensk orðabók*) og Þórdís Úlfarsdóttir (*ISLEX orðabókin*) en þau Kristín Bjarnadóttir og Jón Hilmar Jónsson ræddu efnið almennt og nálguðust það hvort frá sínu sjónarhorni. Þema heftisins er það sama og málþingsins og hér birtast þrjár greinar sem byggðar eru á fyrirlestrum sem þar voru haldnir auk fjórðu greinarinnar sem fjallar um sama efni. Grein Jóns Hilmars Jónssonar, *Að byggja undir íslensk-erlenda orðabók. Forgreining og orðabókarefni*, fjallar um efnisöflun og framsetningu efnisins í ljósi tví- og margræðni uppflettiorðanna og merkingarvensla á milli þeirra með skírskotun til Íslensks orðanets, verkefni sem hann hefur unnið að um árabíl. Kristín Bjarnadóttir ræðir tiltækar aðferðir við söfnun og val uppflettiorða fyrr og nú í grein sem nefnist *Hvert á að sækja orðaforðann í orðabók?* og segir frá nýjum möguleikum sem hafa verið að opnast á síðustu árum með tilkomu rafrænna textasafna og málheilda auk ýmiss konar mál-tæknitóla, þ. á m. til sjálfvirkrar orðtöku úr rafrænum textum. Þórdís Úlfarsdóttir gerir ítarlega grein fyrir gerð og einkennum hinnar norrænu ISLEX-veforðabókar í grein sem nefnist *ISLEX - norræn margmála orðabók* og fjallar þar m.a. um íslensku sem viðfangsmál í því verki. Loks er í þessum hluta grein eftir Rósu Elínu Davíðsdóttur, sem vinnur að doktorsritgerð í orðabókafræði við Háskóla Íslands og Sorbonneháskóla, með titilinn *Hlutverk tvímála orðabóka. Ólíkar not-*

endaþarfir í íslensk-frönsku ljósi. Eins og nafnið gefur til kynna er sjónum beint að hlutverki tvímála orðabóka og þörfum ólíkra notendahópa. Fjallað er um áhrif þessara þátta á efnisöflun og framsetningu efnisins, m.a. eftir því hvort orðabækur eru prentaðar eða miðlað rafrænt. Dæmi eru tekin af orðabókum milli íslensku og frönsku og í bakgrunni eru hugmyndir um gerð nýrrar íslensk-franskrar orðabókar.

Þrjár greinar utan þemans eru birtar í ritinu. Guðrún Kvaran ritar greinina *Orðabókarstörf Konráðs Gíslasonar* þar sem hún rekur vinnu Konráðs við ýmsar orðabækur þótt einungis ein þeirra, *Dönsk orðabók með íslenskum þýðingum* (1851), kæmi út í hans nafni. Aðild hans að öðrum verkum, útgefnum og óútgefnum, er síður þekkt og greinin er því þarft innlegg í sögu íslenskra orðabóka. Hinar greinarnar tvær fjalla báðar um litarheiti í íslensku út frá kenningum mannfræðinganna Berlin og Kay um grunnliti í tungumálum. Grein Susan M. Arthur, *Are Oranges Yellow? Appelsínugulur as a Basic Color Term in Icelandic*, fjallar eins og nafnið bendir til um tilkomu og stöðu litarheitisins *appelsínugulur* í íslensku. Kirsten Wolf segir í sinni grein, *Basic Color Terms in Old Norse-Icelandic: A Quantitative Study*, frá niðurstöðum rannsóknar á tíðni grunnlitarorða í fornu máli og fjallar um samband tíðni og hugmynda um algild innbyrðis vensl slíkra litarorða m.t.t. þess hvort og í hvaða röð þau birtast í tungumálum.

Einn ritdómur er í þessu hefti, umsögn Auðar Lorenzo um *Íslensk-spænska orðabók* (2011), en auk þess hefur ritstjóri tekið saman stuttar ritfregnir um nokkrar nýlegar bækur á sviði tímaritsins. Loks eru sagðar fréttir af áhugaverðum ráðstefnum og málþingum sem haldin verða á næstu mánuðum.

Allar greinar í tímaritinu eru ritryndar og til þess að svo megi vera þarf stór hópur sérfræðinga að leggja hönd á plóginn. Ritrynar njóta nafnleyndar og verða því ekki taldir upp hér en þeim er öllum þakkað fyrir vandaðan yfirlestur og gagnlegar athugasemdir. Allir aðrir sem lagt hafa ritstjóra lið við frágang þessa heftis, einkum meðlimir ritnefndarinnar og samstarfsfólk á stofnuninni, fá þakkir fyrir sitt framlag. Það gildir ekki síst um Bessa Aðalsteinsson sem annast hefur umbrot tímaritsins um árabíl.

Jón Hilmar Jónsson

Að byggja undir íslensk-erlenda orðabók

Forgreining og orðabókarefni

1 Inngangur

Í þessari grein er fengist við að lýsa einkennum og meðferð íslenskra orðabókargagna í tvímála og þá einkum íslensk-erlendum orðabókum.¹ Gerð er stuttlega grein fyrir stöðu íslenskrar orðabókargerðar þessarar tegundar, m.a. með hliðsjón af breyttum aðstæðum með tilkomu netsins og veforðabóka. Lýst er forystuhlutverki viðfangsmálsins og ólíkum aðstæðum eftir því hvort íslenska er í stöðu viðfangsmáls eða markmáls.

Meginefni greinarinnar snýr að íslensku sem viðfangsmáli, út frá efnisföngum, efnisvali og samsetningu flettulistans og því hlutverki hans að mynda umgjörð um orðabókarefnið. Bent er á þær takmarkanir sem eru á flokkun flettulistans og þeirra gagna sem liggja þar að baki og lögð áhersla á gildi þess að gerð hans og uppbygging styðjist við merkingargreiningu og merkingarlega flokkun. Í því sambandi er vísað til *Íslensks orðanets* sem hefur að geyma margbreytileg orðabókargögn með víðtækri greiningu á formgerðarlegum og merkingarlegum venslum þar sem fleiryrtar merkingarbærar einingar hafa sjálfstæða stöðu til jafns við stök orð.

Fjallað er um notagildi stafrænna textasafna og mismunandi grein-

1 Ég þakka ritstjóra og ritrýnendum tímaritsins góðar og gagnlegar athugasemdir og ábendingar við upphaflegan texta greinarinnar.

Orð og tunga 15 (2013), 1–22. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

ingar á þeim með tilliti til lýsingar á notkunareinkennum og notkunarumhverfi orðanna og vali notkunardæma. Þá er vikið að sérkennum íslensks orðabókarefnis þar sem íslenska er markmálið, bent á hversu merkingarþátturinn er gildur við meðferð þess og greiningu og færð rök að því að rækileg forgreining eigi jafnt við hvort sem fengist er við íslensku sem viðfangsmál eða markmál. Loks er fjallað stuttlega um kosti þess að víkja frá hinum orðbundnu venslum málanna og viðhafa hugtakslega nálgun við orðabókarlýsinguna þar sem meira jafnvægi getur orðið með málunum.

2 Orðabók milli mála

Það er eðlilegt metnaðarmál hverju málsamfélagi að bera mál sitt og orðaforða að öðrum tungumálum í mynd orðabóka. Aðstæður til framtaks á því sviði eru vissulega ólíkar frá einu málsamfélagi til annars og mótast m.a. af stærð og krafti málsamfélagsins og hvernig menningar- og félagslegum tengslum þess við önnur málsamfélög er háttað. Landfræðilegt nágreppi, söguleg tengsl og málsögulegur skyldleiki eru virkir hvatar í þessu sambandi en áhrifavald stórra málsamfélaga segir einnig til sín eins og best sést á því hversu orðabækur gagnvart ensku eru almennt fyrirferðarmiklar.

Gerð nýrrar tvímála orðabókar er gjarna a.m.k. öðrum þræði menningarlegt framtak þar sem atbeini beggja málsamfélaga kemur til. Orðabókinni er þá ætlað að styrkja tengsl viðkomandi þjóða, efla viðskipti þeirra, styðja við ferðamennsku o.s.frv. Í öðrum tilvikum skírskotar verkið fremur inn á við, t.d. þegar orðabókin hefur það hlutverk að greiða götu innflytjenda og einstakra innflytjendahópa í nýju samfélagi. Enn önnur verk snúa beinlínis að námi og menntun í viðkomandi tungumáli á ólíkum stigum.

Margt hefur breyst í umhverfi orðabókagerðar og orðabókanotkunar á undanförunum árum. Með tilkomu netsins hefur aðgengi að hvers konar upplýsingum orðið stórum betra en áður og að sama skapi hefur orðið auðveldara að koma upplýsingum á framfæri. Upplýsingar um orð og orðanotkun þurfa hér minni umbúnað en í prentheimi, og þær geta átt gilt erindi í ýmsu formi, t.d. í mynd einfaldra orðalista sem auka má í eftir þörfum og aðstæðum, án þess að um sé að ræða frágengið eða fullbúið orðabókarverk í hefðbundnum skilningi. Notendur netsins njóta góðs af þessu en geta þó stundum átt erfitt með að meta gildi þeirra verka sem ratað er að, hversu vel

er til þeirra vandað og á hverju einstök efnisatriði og upplýsingar eru byggðar.

Þessar aðstæður fela í sér nýjar áskoranir fyrir þá sem standa að gerð og útgáfu orðabóka og annars konar orðabókarefnis. Hinn prentaði búningur heldur að vísu enn velli auk þess sem margar prentaðar orðabækur er jafnframt að finna í rafrænu formi. En netið verður æ virkari upplýsingamiðill fyrir orðabókarefni, sem leitar þar síns forms og framsetningar, án þess að vera bundið af forskriftum prentaðs texta. Einkenni prentaðs orðabókartexta eru þó áfram greinileg enda má gera ráð fyrir að það komi mörgum notendum vel. En jafnframt leitar birtingin meiri sveigjanleika með fjölpættari aðgangi og tengingum innan orðabókartextans og út fyrir hann.

2.1 Íslenska í tvímála samhengi

Nýtt framtak og nýir áfangar í gerð tvímála íslenskra orðabóka mótast af þessum aðstæðum þar sem fjárhagslegar forsendur prentaðra orðabóka veikjast jafnt og þétt og netmiðillinn sækir á. Þörfin fyrir nýjar orðabækur markast þá öðrum þræði af því hversu brýnt þykir að búa lýsingunni umgjörð á netinu hvort sem prentuð orðabók gagnvart viðkomandi tungumáli er fyrir hendi eða ekki.

Tvímála orðabækur um íslensku eru að vonum sundurleitar enda mótaðar af ólíkum aðstæðum á ólíkum tímum. Sumar þeirra eru mjög komnar til ára sinna og hafa enn ekki verið leystar af hólmi en að drjúgum hluta er þó um að ræða nýlegar orðabækur. Að því er varðar tengslin við norræna málsvæðið markar veforðabókin ISLEX, sem tengir íslensku sem viðfangsmál við fimm norræn mál (dönsku, norskt bókmál, nýnorsku, sænsku og færeysku) vissulega tímamót, og gerð hennar opnar vonandi leiðir að sambærilegum samanburði við önnur markmál í mynd tví- eða margmála orðabóka á vefnum. Sem veforðabók getur hún líka betur en fyrri prentaðar orðabækur haft gildi við ólík notkunartilefni og sameinað þarfir íslenskra og erlendra notenda.

Þær orðabækur gagnvart norrænum málum þar sem íslenska er í hlutverki markmálsins eru harla ólíkar, og í þeim flokki er skýr þörf á umbótum og endurnýjun sem m.a. felur í sér notendaaðgang á vefnum. Yngst er *Dönsk-íslensk orðabók*, að stofni til frá 1992 og aukin í nýrri útgáfu 2004, en orðabækur gagnvart norsku og sænsku eru ívið eldri (*Svensk-isländsk ordbok* (1982), *Norsk-isländsk ordbok* (1987)).

Þegar litið er út fyrir norræna málsvæðið er staðan sem vænta

má í ýmsum tilvikum veikari, sérstaklega að því er varðar íslensku sem viðfangsmál. Gagnvart þýsku er augljós þörf á nýrri og vandaðri orðabók þótt þær íslensk-þýsku orðabækur sem tiltækar eru geti greitt götu margra notenda (*Íslensk-þýsk orðabók* (1993) og *Þýsk-íslensk/íslensk þýsk orðabók* (2004)). Gagnvart frönsku er staðan þó sýnu verri þar sem eina íslensk-franska orðabókin, *Íslenzk-frönsk orðabók* (1950), er orðin meira en hálfrar aldar gömul. Hér háttar hins vegar betur til þar sem íslenska er markmálið, því *Frönsk-íslensk orðabók* kom út árið 1995 og *Þýsk-íslensk orðabók* árið 2008. Gagnvart ensku er einnig þörf á endurnýjun þótt þar sé staðan betri. *Ensk-íslensk orðabók* (1984) er sannarlega efnismikil og var mikilsvert framtak á sínum tíma. Forlagið fylgdi þeirri útgáfu eftir með *Ensk-íslenskri skólaorðabók* (1986) sem síðan kom út aukin og endurbætt undir heitinu *Ensk-íslenska orðabókin* (2006). Íslenska er svo viðfangsmálið í *Íslensk-enskri orðabók* Sverris Hólmarssonar, Christophers Sanders og Johns Tucker, sem fyrst kom út árið 1989 en að nýju aukin og endurbætt í ritstjórn Christophers Sanders árið 2009. Ítölsku og spænsku hafa verið gerð skil með samstæðum orðabókum í báðar áttir (*Íslensk-ítölsk orðabók* (1994) og *Ítölsk-íslensk orðabók* (1999), *Spænsk-íslensk orðabók* (2007) og *Íslensk-spænsk orðabók* (2011)). Flestar þessara orðabóka eru jafnframt tiltækar sem veforðabækur í vefbókasafninu *Snöru*. Meðal tvímála orðabóka um íslensku ber svo sérstaklega að nefna hinar efnismiklu orðabækur gagnvart rússnesku, *Íslenzk-rússneska orðabók* (1962) og *Rússnesk-íslenska orðabók* (Helgi Haraldsson 1996) sem um margt hafa líkt yfirbragð þótt sú síðarnefnda sé enn efnismeiri og margbrotnari að gerð. Í báðum þessum bókum er orðalýsingin að ýmsu leyti rækilegri en notendur eiga að venjast í tvímála orðabókum.

2.2 Hlutverk í tvímála orðabók: viðfangsmál eða markmál

Hér verða einstakar tvímála orðabækur um íslensku ekki tíundaðar frekar þótt nefna megi ýmis markverð verk á því sviði. Þess í stað verður athyglinni beint að þeim aðstæðum og forsendum sem gerð slíkra orðabóka býr við, sérstaklega að því er varðar grundvallargögn, val, afmörkun og framsetningu orðabókarefnisins og aðgang notenda að orðalýsingunni og einstökum þáttum hennar. Þar er skýr munur á eftir því hvoru hlutverkinu íslenska gegnir, hvort hún er viðfangsmálið og þar með í forgrunni orðabókartextans (íslensk-erlendar orðabækur),

eða markmál sem bregður upp samsvörunum við erlent viðfangsmál orðabókarinnar hverju sinni (erlend-íslenskar orðabækur).

2.2.1 Viðfangsmálið í forgrunni

Viðfangsmálið og efnisþættir þess móta að jafnaði umgjörð tvímála orðabókar og afmarka umfang hennar og orðafjölda. Þar myndar flettulistinn uppistöðuna sem öðrum efnisþáttum er skipað undir. Þegar íslenska er í þessu hlutverki kemur það í hennar hlut að tefla fram þessum uppistöðueiningum, sem innan viðeigandi marka eiga að birta samfellda mynd af íslenskum orðaforða og orðanotkun. Þar með verður val flettiorða eðlilega í brennidepli á fyrstu stigum orðabókarverksins.

Samsetning flettulistans getur átt sér stað með ýmsum hætti og margar einstakar ákvarðanir ráðast fyrst og fremst af mati þess, höfundar eða ritstjóra, sem um vélar og ábyrgðina ber hverju sinni. Hinn áþreifanlegi vandi felst oftast en ekki í því hversu erfitt er að beita samræmdum mælikvarða á flettugildi orða, meta hvort tiltekið orð eigi að fá inni í listanum (og fái þá jafnframt jafnheiti eða annars konar skýringu á markmálinu) eða hafi einhver þau einkenni sem réttlæta að litið sé framhjá því. Þar koma til ólík greinimörk og tillit til ólíkra notendahópa. Það snertir m.a. mat á hlut samsetninga og virkri orðmyndun, viðhorf til málstöðlunar og málræktar og hvort orðabókarlýsingin á sem mest að einskorðast við samtímamál.

Annar og raunar stærri vandi, sem í vissum skilningi er óáþreifanlegur, lýtur að því hversu takmarkaðar og brotakenndar birtingarmyndir standa til boða af íslenskum orðaforða, samsetningu hans og notkunareinkennum. Við þær aðstæður getur þótt nærtækt að njóta eftir föngum góðs af þeim ákvörðunum sem fyrir liggja í öðrum hliðstæðum orðabókum og spara að sama skapi fyrirhöfn við sjálfstæða efnisöflun. Um þennan vanda og viðbrögð við honum verður nánar fjallað hér á eftir, einkum í 3. og 4. kafla.

2.2.2 Markmálið: samsvaranir við annað mál

Þegar íslenska er í hlutverki markmáls eru aðstæður í grundvallaratriðum aðrar. Þar þarf að finna íslenskar samsvaranir við einingar annars tungumáls sem mótar flettiorðavalið á sínum forsendum. Sá íslenski orðaforði sem ratar inn í það umhverfi ber svip af því. Ýmis gild orð með skýru flettugildi koma jafnvel ekki við sögu á meðan

önnur og óvirkari orð eru dregin fram eða mynduð eru ný orð til að ná fram viðeigandi samsvörun í mynd íslensks jafnheitis. Annars er eðli málsins samkvæmt erfitt að fá heillega yfirsýn um orðaforðann og orðafarið markmálmegin, eins dreift og það er um orðabókartextann. Hér er hlutverk þess sem á heldur að hnitmiða jafnheitavalið eins og kostur er. Þar getur stundum virst fátt um vænlegar lausnir, í öðrum tilvikum virðast ólíkar leiðir í boði, a.m.k. við fyrstu sýn.

Einnig við þessar aðstæður er hætt við að erfitt reynist að framkalla heillegar myndir af íslenskum orðaforða og einstökum hlutum hans til stuðnings jafnheitavalinu. Að því leyti er enn erfiðara um vik að það orð eða orðafar sem rata skal að á sér engan formlegan íslenskan búning fyrir fram heldur tengist það og kviknar af merkingarlegum eigindum orða sem fengist er við í erlenda viðfangsmálinu. Íslenska samsvörun sem fyrst kemur upp í hugann þarf að máta við önnur hugsanleg jafnheiti á sama merkingarsviði en þar er yfirsýn oftast en ekki takmörkuð.

3 Efnisföng, efnisval og flettulisti

Í tvímála orðabók er teflt saman orðaforða tveggja tungumála með skírskotun til markhópa báðum megin þótt með ólíkum hætti sé. Markhópurinn markmálmegin nálgast lýsinguna fyrst og fremst til skilnings á einstökum orðum og orðasamböndum viðfangsmálsins og til glöggvunar á orðanotkun í því. Markhópurinn sem hefur viðfangsmálið að móðurmáli eða er handgenginn því leitar á hinn bóginn samsvarana við einingar þess (flettiorð og orðasambönd) í markmálinu.

Í íslensk-erlendri orðabók er þannig jöfnum höndum hugsað til þarfa íslenskra og erlendra notenda. Í fræðilegri umræðu um tvímála orðabækur hefur mjög verið fjallað um mismunandi notkunarhlutverk og vandkvæði á að þau fari saman í einu orðabókarverki (sbr. m.a. Svensén 2004:17–23). Þannig geti orðabók sem leggur áherslu á að greiða fyrir skilningi á merkingu orða í viðfangsmálinu ekki að sama skapi þjónað þeim sem leita upplýsinga um orðanotkun (notkunarþætti jafnheita) í markmálinu. Samkvæmt því hafa margar tvímála orðabækur verið samdar með býsna sérhæft hlutverk í huga. Slík aðgreining tekur reyndar mið af prentuðum orðabókum en í orðabókum á netinu er mun hægara að koma til móts við mismunandi þarfir ólíkra notendahópa, með breytileika í efnisröðun, beinni leit

að efnisþáttum innan orðsgreina og innbyrðis tengingum. Hvort sem hægt er að líta svo á að gagnvart íslensku sé aðeins rúm fyrir eina tvímála orðabók í hvora átt um hvert málapar (t.d. eina íslensk-enska og eina ensk-íslenska, sbr. Sanders 2005:53) styrkir þetta mjög netið sem vænlegan miðil þegar hugsað er til nýrra tvímála íslenskra orðabóka.

3.1 Flettulistinn sem umgjörð um orðabókarefnið

Gagnvart íslensku í íslensk-erlendri orðabók er viðfangsefnið í upphafi að afla efniviðar til að byggja upp íslenskan flettulista og greina notkunarþætti flettiorðanna, aðgreina merkingarbrigði, tilgreina orðasambönd og lýsa einkennandi orðanotkun. Í þessu ferli er afmörkun flettulistans gjarna í fyrirrúmi, þar sem fyrst er úrskurðað um flettugildi orða í því safni orða sem efnisöflunin leggur til, með tilliti til áætlaðrar stærðar verksins og hlutverksins sem því er ætlað. Síðan er lýsing hvernar flettu tiltölulega sjálfstætt og afmarkað verkefni, þar sem gera þarf sérstaka atrennu að efnisöflun, velja úr tiltæku efni og setja lýsingunni hæfilegar skorður. Önnur aðkoma er að gefa orðanotkuninni, og þá sérstaklega merkingarbærum orðasamböndum, meiri gaum þegar við fyrstu afmörkun samhliða flettuvalinu.

Hvernig sem afmörkuninni er háttað kallar hún á viðeigandi verklag við lýsingu orðanna sem mynda flettulistann. Hætt er við að erfitt reynist að fá æskilega yfirsýn um samsetningu flettulistans svo að hægt sé að nálgast verkið í skipulegum áföngum. Þar togast m.a. á hagræðið að því að láta stafrófsröðina vísa veginn og meta út frá því hvernig verkinu miðar, fást við hvern og einn orðflokk sérstaklega og treysta með því innbyrðis samræmi eða fylgja eftir lýsingu tiltekinna merkingarflokka.

3.2 Þörf á merkingarflokkun

Almenn flokkun og greining flettulista og annarra orðalista takmarkast að mestu við orðflokkaskiptingu. Að auki dregur stafrófsröðin fram orðmyndunarleg vensl og þau vensl geta komið frekar í ljós ef leitað er tiltekinna bakliða samsettra orða, t.d. fyrirferðarmikilla nafnorðaliða eins og *-maður* eða lýsingarorða með liðum eins og *-látur* og *-laus*. Merkingarlegir flokkunarþættir eru hins vegar yfirleitt ekki fyrir hendi og þá ekki heldur gagnvart fleiryrtum merkingarbærum einingum sem gerð eru skil innan flettugreina.

Þörfin fyrir merkingarlega flokkun er þó næsta augljós enda er sam-
anburður viðfangsmáls og markmáls merkingarlegs eðlis. Vandinn er
hins vegar sá að ekki dugir að staðnæmast við flettimyndirnar ein-
ar heldur verður jafnframt að greina á milli merkingarbrigða tví- og
margræðra orða. Miðað við hefðbundið lýsingarferli eru því ekki for-
sendur fyrir merkingarlegri flokkun orðabókarefnisins fyrr en allri
margræðni hafa verið gerð skil.

Af þessu leiðir að merkingarleg flokkun verður ekki það viðfangs-
efni í orðabókarferlinu sem hún í raun þyrfti að vera, og reynslan úr
einu orðabókarverki nýtist að takmörkuðu leyti í öðru sambærilegu
verki. Gildi merkingarflokkunar snýr bæði að viðfangsmálinu sem
slíku og að samhengi þess við markmálið. Ef slík flokkun tekur til
tiltæks orða- og orðasambandasafns í heild sinni án tillits til síðari
takmörkunar og endanlegs efnisvals getur hún gegnt virku hlutverki
við sjálft flettuvalið, og sá orðaforði sem út af stendur er þá áfram í
greiningartengslum við flettulistann. Þegar kemur svo að lýsingu og
greiningu flettnanna skiptir miklu að hafa yfirsýn um merkingarskyld
orð og sambönd, svo að ná megi fram samfellu í greiningunni með
hliðsjón af skyldu orðafari. En gagnsemin er jafnvel enn áþreifanlegri
fyrir markmálið, því flokkunin dregur fram það orðafar í markmál-
inu sem til álita kemur að tilgreina sem jafnheiti og stuðlar þannig að
hnitmiðuðu jafnheitavali og markvissum skýringum.

3.3 Merkingarflokkun í *Íslensku orðaneti*

Merkingarflokkun af þessu tagi hefur að nokkru leyti sannað gildi sitt
við greiningu orðabókarefnisins í ISLEX (sjá Þórdís Úlfarsdóttir (2013
í þessu hefti)) en þar er byggt á víðtækari og margþættari flokkun á
mun stærra safni orða og orðasambanda í *Íslensku orðaneti* (sjá Jón
Hilmar Jónsson 2012). Gildi orðanetsins í þessu sambandi er að miklu
leyti fólgið í því hvernig flettimyndum þess og flettuskipan er háttað.
Annars vegar eru flettunar merkingarlega einræðar þannig að hvert
merkingarbrigði er sjálfstæð eining. Hins vegar rúmar flettulistinn
auk stakra orða (einyrtra flettna) mikinn fjölda fleiryrrtra flettna þegar
um er að ræða merkingarbær orðasambönd. Þannig má greiðlega
tengja greiningu orðanetsins við flettiorð, einstök merkingarbrigði og
merkingarbær orðasambönd innan tiltekins orðabókartexta.

Auk merkingarflokka í mynd lýsandi hugtakaheita (nafnorða
eins og ÁHUGI, ÁHUGALEYSI, ÁHUGAMÁL, ÁHYGGJUR, ÁHYGGJULEYSI,
ÁHÆTTA) kemur merkingargreining orðanetsins fram í samheitum og

andheitum. Innan merkingarflokksins ÁHYGGJUR eru einyrtaflettur m.a. *áhyggjur, búksorgir, hugarvöl og sút* og meðal tvíyrtra flettna eru *hafa áhyggjur, bera ugg í brjósti, vera uggandi um sinn hag, vera í öngum sínum, <honum, henni> vex <þetta> í augum og <þetta> stendur <mér> fyrir svefni*. Flettan *<þetta> stendur <mér> fyrir svefni* er auk þess í samheitavenslum við flettuna *<þetta> heldur vöku fyrir <mér>* og sams konar vensl eru á milli flettnanna *láta sér <þetta> í augum vaxa og mikla <vandann> fyrir sér*.

Þannig getur *Íslenskt orðanet* stutt merkingarflokkun í greiningar- og vinnsluferli íslensk-erlendra orðabóka þótt gera verði ráð fyrir að slík flokkun geti seint orðið tæmandi. Í máltækniverkefninu *Íslenskur merkingarbrunnur* (sjá Anna B. Nikulásdóttir 2012) er einnig fyrir hendi viðtæk merkingarflokkun á íslenskum orðaforða sem hagnýta má í sama skyni en þar er beitt öðrum og vélrænni aðferðum við greininguna. Gögnin sem þessi verkefni leggja til geta gegnt virku hlutverki þegar byggt er upp orða- og flettusafn íslensk-erlendrar orðabókar og auðveldað efnisafmörkun sem hæfir því hlutverki sem orðabókinni er ætlað að gegna.

4 Formgerð og framsetning merkingarbærra orðasambanda

Hér að framan var minnst á orðflokkgreiningu sem helsta greiningarþátt orðabókarefnis samkvæmt rótgróinni hefð. Þótt orðflokkgreining eigi sér langa sögu hefur hún sjaldnast gegnt beinu flokkunarhlutverki og segja má að takmörkuð áhersla hafi verið lögð á sérstöðu einstakra orðflokka í almennum orðabókum. Í stafrænu umhverfi eru breyttar forsendur að þessu leyti, hægt er að skilyrða efnisflokkun og orðaleit við einstaka orðflokka og hægara er að samræma meðferð orðflokka í orðabókartextanum. En hefðbundin orðflokkgreining nær aðeins til stakra orða og með því verða fleiryrtar merkingarbærar einingar út undan og eiga sér ekki sambærilegan greiningarþátt.

Í *Íslensku orðaneti* er viðhöfð málfræðileg mörkun á fleiryrtum flettum sem gefur færi á formgerðarlegri flokkun þeirra (sjá Jón Hilmar Jónsson 2012:46-49). Mörkunin tekur að nokkru leyti mið af hinu orðabókarlega umhverfi og er í vissum atriðum einfaldari en venjubundin málfræðileg textamörkun. Auk orðflokkgreiningar er megináherslan á fallmörkun, lexíkalskan breytileika í orðasamböndum og innbyrðis

vensl setningarliða. Markið so <no-dg> fs no-a (sögn með eftirfarandi nafnorði í þágufalli með greini, að viðbættri forsetningu með nafnorði í þolfalli) sameinar t.d. flettur *koma* <húsinu> í *stand*, *stofna* <samkomulaginu> í *hættu*, *kasta* <hlutleysinu> fyrir *borð* og *skjóta* <ákvörðuninni> á *frest*, þar sem oddklofarnir afmarka lexíkalskan breytileika. Á sama hátt tengir markið <no-ag> so saman flettur *<gluggann> hrímar*, *<bátinn> fyllir*, *<sjóinn> gárar* og *<vindinn> hægir*. Formgerðarflokkunin kallast svo eftir atvikum á við merkingarlega greiningu og flokkun á fleiryrtum flettum og styður hana reyndar gjarna þar sem samfella í formgerð á sér oft samsvörun í merkingarlegum skyldleika.

Mörkunin byggist á því að fleiryrtar flettur í orðanetinu eru settar fram á staðlaðan hátt, m.a. með því að afmarka og auðkenna orðbundinn breytileika. Þessa meðferð fleiryrttra merkingarbærra eininga má nýta utan orðanetsins til greiningar og flokkunar á orðabókarefni svo að betri yfirsýn gefist um orðasambönd, fyrirferð þeirra og stöðu innan orðabókartextans. Slíkt notagildi er í sjálfu sér óháð því hvernig merkingarbærum orðasamböndum er fyrir komið í textanum og hvaða snið er haft á þeim. Hliðsjón af orðanetinu og samanburður við almennar orðabækur getur hins vegar sýnt hversu þröngt er um orðasambönd sem undirskipaðar einingar innan orðabókartexta og hversu erfitt er að gera þau sýnileg á sjálfstæðan hátt.

Atviksliðasambönd af ýmsu tagi eru skýrt dæmi um fleiryrtar merkingarbærar einingar sem full þörf er á að séu í sjónmáli innan viðfangsmálsins í tvímála orðabókartexta. Upphafseiningar slíkra sambanda eru gjarna forsetningar, atviksorð eða samtengingar, en þeim orðflokkum er jafnan gefið takmarkað rúm þar sem hið málfræðilega hlutverk viðkomandi orðs er í fyrirrúmi umfram einstök sambönd. Slík sambönd skila sér hins vegar ekki sem skyldi ef þeim er ætlað sæti undir gildari orðum, t.d. nafnorðum eða fornöfnum, og þar getur reynt erfitt að rata að einstökum samböndum í efnismiklum flettugreinum.

Á vinnslustigi getur einnig reynt torvelt að fá æskilega yfirsýn um margs kyns sagnasambönd með smáorði og sama má segja um orðtök í mynd sagnliða þar sem oft ber á breytileika í orðum og orðmyndum. Þar getur mörkun fleiryrttra eininga í stærra gagnasafni komið að góðum notum.

5 Notkunareinkenni og notkunardæmi

Forgreining orðabókargagna af því tagi sem lýst hefur verið hér að framan á fyrst og fremst að hafa almennt gildi og vera óháð því hvernig efnisskipan er háttáð í einstökum orðabókum sem notfæra sér gögnin. Almennir flokkunarþættir, málfræðilegir sem merkingarlegir, veita aðhald og yfirsýn, hvort sem efnisskipan í endanlegum texta er með hefðbundnu sniði eða hún er látin mótast af greiningunni í ríkari mæli. En greiningin stuðlar að frjálssari efnisskipan þar sem merkingarleg einkenni geta fengið aukið vægi og fleiryrt sambönd haft meira sjálfstæði en hefð er fyrir eins og greinarhöfundur hefur áður lýst (sjá Jón Hilmar Jónsson 2005).

Hvernig sem efnisskipaninni er háttáð er lýsing viðfangsmálsins að drjúgum hluta fólgin í því að draga fram einkennandi notkun orða og orðasambanda, m.a. með beinum dæmum. Fyrirferð notkunardæma getur verið afar breytileg og í raun er það á valdi ritstjóra hverju sinni hve miklu hlutverki þau eru látin gegna og hvernig þau eru sett fram. Í texta tvímála orðabókar er staða notkunardæma að því leyti önnur og sjálfstæðari en í einmála orðabók að dæmin styðjast ekki við undanfarandi merkingarskýringu á sama máli og skerpa þá mynd sem þar kemur fram, heldur staðfesta þau merkingarlega samsvörun í mynd erlends jafnheitis eða jafnheita. Notkunardæmin eiga því fyrst og fremst erindi við þá notendur sem nálgast viðfangsmálið sem erlent mál og hafa um leið vald á markmálinu.

Þessar aðstæður fela í sér ákveðinn vanda að því leyti að annars vegar þarf að velja og mynda notkunardæmin með skýrri yfirsýn um viðfangsmálið, íslenska orðanotkun, breytileika hennar og blæbrigði. Hins vegar kemur til þörf erlendra notenda fyrir að fá fram tiltekin megineinkenni í orðanotkun, sérstaklega með tilliti til þess að geta tileinkað sér málið og beitt því við raunverulegar aðstæður.

5.1 Form og framsetning notkunardæma

Form og framsetning notkunardæma getur verið með ólíkum hætti eftir því hvaða hlutverki þeim er ætlað að gegna. Þegar áherslan er á setningargerð er í norrænni orðabókarhefð gerður greinarmunur á „lifandi“ og „dauðum“ dæmum („levande og döda språkprov“, sjá Svensén 2004:193–195). Hin fyrrnefndu standa nær beinu notkunarsamhengi, með nafnliðum í mynd nafnorða eða persónufornafna

og sögnum í persónuhætti en hin síðarnefndu birta eins konar notkunarmunstur þar sem notkunarsamhengið er tilgreint á knappan og formfastan hátt, sagnir í nafnhætti og nafnliðir í mynd óákveðins fornafns. Í íslenskum orðabókum eru slík munsturdæmi einkum bundin skammstöfunum fornafnsins *einhver*, t.d. *hitta e-n*, *mæta e-m/e-u*, *gefa e-m e-ð*. Munsturdæmi draga fram tiltekin megineinkenni eins og fallstjórn en þau eru ekki til þess fallin að sýna merkingareinkenni og merkingarlega afmörkun nálægra orða eins og frumlags eða andlags sagna og því getur þótt ástæða til að fylgja þeim eftir með frjálsari dæmum. Þriðji kosturinn er að sameina þessi dæmahlutverk með því að tilgreina dæmigerða fulltrúa fylgdarorða með skýrum málfræðieinkennum í formföstum dæmum eins og m.a. er gert í *Stóru orðabókinni um íslenska málnotkun* (2005) og síðan í *Íslensku orðaneti*. Notkun persónufornafnsins í stað fornafnsins *einhver* gegnir þá því hlutverki að ljá dæmunum skýrari svip beinnar málnotkunar: *hitta* <*hann, hana*>, *mæta* <*honum, henni*>, *mæta* <*andstöðu, mótspyrnu; andstreymi*>, *gefa* <*honum, henni*> <*gjöf; mat; leyfi*>. Frá öðru sjónarhorni getur dæmigerð orðanotkun birst í virkum og áberandi orðastæðum, þar sem flettiorðið kemur fram með einkennandi fylgdarorði: *rennandi blautur* (undir *blautur*), *brimið svarrar* (undir *brim*), *detta kylliflatur* (undir *detta*).

5.2 Samhengið við merkingarbær orðasambönd

Hvaða framsetning sem valin er á notkunardæmum í íslensk-erlendri orðabók hvílir lýsing íslensku flettiorðanna að verulegu leyti á því dæmasafni sem þau eru látin styðjast við. En dæmapörfin snertir ekki aðeins flettiorðin sem slík og breytilega notkun þeirra heldur getur hún einnig átt við um margs konar sjálfstæð og föst orðasambönd sem skipað er undir flettiorðin. Þá getur vandinn verið sá hvernig megi takmarka og einfalda lýsingu einstakra flettugreina og þar með dæmasafnið svo að notendur missi ekki yfirsýn um efnisatriðin í ofhlöðnum orðabókartexta.

Sú áskorun reynist að vonum erfið í prentuðum orðabókartexta, eins og margar íslensk-erlendar orðabækur vitna um. Raunin er oft sú að föstum merkingarbærum samböndum og almennari notkunardæmum orðanna er skipað saman, utan við aðgreiningu merkingarbrigða sem gerð er grein fyrir með ólíkum jafnheitum, og innbyrðis röð gjarna látin ráðast af forsetningum eða atviksorðum sem fram koma í samböndunum (sjá Jón Hilmar Jónsson 1995). Þar með rofn-

ar samhengi merkingareinkenna orðsins við setningargerðina og merkingarbær orðasambönd skera sig ekki úr hópi annarra sambanda orðsins sem sjálfstæðar einingar. Slík tilhögun er í alla staði óheppileg en þó sérstaklega fyrir þá notendur sem leita leiðsagnar um málbeitingu á íslensku.

5.3 Greiðari leið að dæmasöfnum

Í veforðabók má gefa notkunardæmum mun meira svigrúm og um leið skýrara og margþættara hlutverk en tók eru á í prentaðri orðabók. Þau þurfa t.d. ekki í öllum tilvikum að birtast á yfirborði orðabókartextans heldur verið framkallanleg ef notandinn kys að skerpa tiltekna umsögn og notkunareinkenni. Aðstæður leyfa einnig að merkingarbær orðasambönd standi sjálfstætt og komi eftir atvikum fram með sínum notkunardæmum. Þá verður sá kostur æ nærtækari að leita dæma utan við orðabókartextann með því að svipast um í flokkuðu safni textadæma viðkomandi orðs í mynd orðstöðulykils eða markaðrar málheildar (e. *tagged corpus*). Notendum dönsku vefsíðunnar dsl.dk/sprog (sem *Det danske sprog- og litteraturselskab* stendur að) gefst t.d. aðgangur að stórrí danskri málheild, *KorpusDK*, samhliða aðgangi að stórum dönskum orðabókarverkum. Þangað geta notendur orðabókanna sótt viðbótardæmi um einstök orð og virt nánar fyrir sér stöðu orðanna í setningarlegu samhengi.

Að því er orðanotkun og notkunardæmi varðar má nú auðveldlega leita fanga með beinni leit í tölvutækum textum, hvort sem það er gert í tengslum við orðabókarnotkun eða óháð þeim aðstæðum. Stundum á best við að leita fanga í ógreindu textasafni, t.d. hinu geysistöra safni *Tímarit.is* sem tekur til íslenskra blaða og tímarita frá upphafi til dagsins í dag. Í öðrum tilvikum hentar betur að nýta sér greiningu sem fram hefur farið á viðkomandi textum. *Íslenskur orðasjóður* (2012) er textasafn af því tagi, sótt í íslenskar vefsíður, þar sem setningarlegt umhverfi orðanna hefur verið greint og notendur geta nýtt sér þá greiningu, t.d. svipast um í hópi áberandi vinstri eða hægri nággranna orðsins og ratað þaðan að beinum textadæmum. *Íslenskt textasafn* Stofnunar Árna Magnússonar í íslenskum fræðum hefur að geyma fjölbreytta íslenska texta þar sem raða má textadæmum eftir orðmynd leitarorðsins eða í stafrófsröð eftir undanfarandi eða eftirfarandi orði. *Mörkuð íslensk málheild* sameinar svo safn íslenskra texta frá árabílinu 2000–2009 þar sem orðin hafa fengið rækilega málfræðilega greiningu og leita má dæma um einstök orð og orðmyndir í tengslum við

málfræðileg og setningarleg einkenni eða draga fram slík einkenni óháð því hvaða orð eiga í hlut.²

Í fyrrgreindum textasöfnum má ganga að forgreiningu á ólíku stigi og hafa hana til stuðnings og hliðsjónar við val og framsetningu notkunardæma í orðabókartexta. Slík greining stuðlar að markvissara dæmavali en hún hefur vitaskuld einnig verulegt gildi gagnvart öðrum þáttum orðabókartextans, svo sem skilum og breytileika í merkingu, áberandi orðastæðum og ekki síst einkennandi setningargerð.

5.4 Orðastæður og forgreining notkunarsambanda

Gagnvart orðabókarnotum er þörf á greiningu sem færir notkunarsambönd nær búningi eiginlegrar orðabókar, dregur saman sambönd sömu formgerðar, sameinar merkingarskyldar orðastæður og gerir merkingarbær orðasambönd sjálfstæð og leitarbær. Í gagnagrunni *Íslensks orðanets* er slík greining fyrir hendi í verulegum mæli eins og að nokkru leyti hefur verið gerð grein fyrir hér að framan. Í flettuskipan orðanetsins birtist hún m.a. í tvíliða lýsingarorðaflettum þar sem nefnimynd lýsingarorðsins er augin með dæmigerðum fulltrúa eða fulltrúum nafnorðs sem lýsingarorðið á við: **sterkur [band, þráður], sterkur [ljós, birta], sterkur [vindur], sterkur [tilfinning]**. Þannig eru merkingarblæbrigði lýsingarorðsins aðgreind, og að baki einstökum flettum er að finna sambönd (orðastæður, e. *collocations*) með breytilegum nafnorðum. Þar sameinar t.d. flettan **sterkur [tilfinning]** sambönd eins og *sterk tilfinning*, *sterk kennd*, *sterk samkennd*, *sterkar ástriður*, *sterk innlifun* og *sterk löngun*. Nafnorðsþáttur flettunnar getur svo haft sameinandi hlutverk gagnvart lýsingarorðum í sömu stöðu: **frískur [vindur], snarpur [vindur], stífur [vindur]**. Á sama hátt geta fleiryrtar sagnaflettur sameinað orðastæður með merkingarskyldum nafnorðum, eins og flettan **borga <gjaldið, skuldina, launin>** þar sem m.a. er að finna sambönd eins og *borga burðareyri*, *borga brúartoll*, *borga inngangseyri*, *borga félagsgjald* og *borga meðlag*. Það svið má svo enn víkka út með því að líta til nafnorða sem fram koma undir flettunni **greiða <gjaldið, skuldina>**. Þau eru sem vænta má að nokkru leyti hin sömu en þar koma einnig fram sambönd eins og *greiða burðargjald*, *greiða árstíllag*, *greiða lóðargjald*, *greiða sekt*, *greiða tekjuskatt*. Þannig getur merkingarflokkun flettnanna fært saman merk-

2 Verkefnið er enn í vinnslu en leitaradgangur er nú þegar að hluta efnisins á vefsíðunni <http://mim.hi.is/> (rúmlega 17 milljón lesmálsorðum af þeim 25 milljónum sem fyrirhugað er að verði í málheildinni fullbúinni).

ingarskyld orð og veitt yfirsýn um sambönd og orðastæður á víðu merkingarsviði.

6 Íslenska sem markmál

Þau orðabókargögn og sú gagnagreining sem hér hefur verið lýst hefur ekki einhliða gildi fyrir íslensku sem viðfangsmál í tvímála orðabókum. Raunar miðast hún ekki síður við einmála orðabókargerð þar sem gerð er grein fyrir íslenskum orðaforða og íslenskri orðanotkun á íslensku. Báðar orðabókartegundirnar hafa þörf fyrir greiðan aðgang að umfangsmiklu gagnaefni sem þær síðan velja úr og skipa á ólíkan hátt eftir því hvernig viðkomandi orðabókarverk er hugsað og afmarkað. En íslensku er einnig lýst á sinn hátt þegar henni er stillt upp gagnvart öðru viðfangsmáli í erlend-íslenskri orðabók. Eins og fyrr var drepið á eru aðstæður gjörólíkar í tvímála orðabók af því tagi, þar sem gagnapörfin miðast ekki við sjálfstæða efnisskipan heldur ræðst íslenska orðabókarefnið og meðferð þess af því hvernig háttar til í efnisskipan og framsetningu viðfangsmáls orðabókarinnar. Meginverkefnið er í því fólgið að velja viðeigandi íslensk jafnheiti við þau flettiorð, merkingarbrigði og orðasambönd sem þar koma fram. Ákvarðanir í því efni þurfa að geta stuðst við viðeigandi gögn um íslenskan orðaforða og orðanotkun og þar kemur greining af því tagi sem hér hefur verið fjallað um m.a. að notum.

6.1 Yfirsýn um íslensk jafnheiti

Íslenskt orðabókarefni í stöðu markmáls er að sjálfsögðu ósamfelldara en þegar um er ræða sjálfstæða efnisskipan gagnvart tilteknu markmáli. Í prentuðum orðabókartexta fæst engin heildstæð yfirsýn um íslenska orðafarið, hvorki orðaforðann sem slíkan og einkenni hans né stöðu einstakra jafnheita, að hvaða marki þau koma fram gagnvart ólíkum orðum og samböndum viðfangsmálsins, hvar brugðið er á það ráð að umorða tiltekna merkingu í stað þess að tilgreina jafnheiti o.s.frv. Þegar orðabókartextinn er í stafrænu formi getur fengist skýrari mynd en heildarmyndin skilar sér ekki nema með sérstakri gagnavinnslu þar sem íslensku jafnheitunum er raðað innbyrðis og þeim stillt upp andspænis þeim einingum viðfangsmálsins sem þau eru látin samsvara í orðabókartextanum.

Fullyrða má að sá íslenski orðaforði sem jafnheitin mynda og í ljós

kæmi við gagnagreiningu erlend-íslenskra orðabóka sé með nokkuð öðrum svip en sá sem valinn er þegar íslenska er viðfangsmálið. Þegar mótuð jafnheiti eru ekki tiltæk getur þurft að bregðast við með nýmyndunum af einhverju tagi, tókupýðingum eða nýgervingum. Glíman við erlenda málið getur vakið áleitnar spurningar um aðild og viðurkenningu aðkomuorða og jafnheitavalið þannig fengið málpólítískt gildi. Orðmyndunarleg sérkenni hvors máls um sig torvelda oft jafnheitavalið og við þær aðstæður er mikilvægt að geta haft hliðsjón af fyrri lausnum og eftir atvikum fetað þar troðnar slóðir. Allt þetta undirstrikar gildi þess að að geta betur en nú gefst kostur á gengið að og hagnýtt það íslenska orðabókarefni sem erlend-íslenskar orðabækur leggja til. Þá er ekki einungis horft til orðabóka þeirrar tegundar því gagnlegt er að bera þann orðaforða sem hér er á ferðinni að orðaforða íslensk-erlendrar orðabókar. Þótt svipurinn sé ólíkur má gera ráð fyrir að meðal jafnheitanna komi fram orð sem ástæða sé til að bæta við flettulistann hinum megin, svo sem gegnsæjar mótaðar samsetningar sem ekki hafa augljóst flettugildi frá íslensku sjónarhorni en verða gildari að því leyti í ljósi erlendrar samsvörunar.

6.2 Greining á merkingarvenslum

Vandað og hnitmiðað val á jafnheitum er meginatriði við gerð tvímála orðabókar. Til þess að tryggja það sem best þarf að gefast yfirsýn um orðafar sem til greina kemur og það merkingarsvið sem um er að ræða hverju sinni. Í því sambandi getur *Íslensk samheitaorðabók* (2012) vísað veginn til merkingarskyldra orða en víðtækari og fjölþættari merkingarflokkun af því tagi sem áður var lýst og *Íslenskt orðanet* býr yfir er einnig mikilvæg stoð. Að auki gefur orðanetið færi á að meta stöðu einstakra orða með tilliti til merkingarskyldleika við önnur orð á grundvelli setningarlegra tengsla þeirra innbyrðis. Þar er fyrst og fremst um að ræða þann vitnisburð sem lesa má úr tengslum orðapara með tengiorðinu og meðal nafnorða og lýsingarorða (*fátækt og allsleysi, allsleysi og nauð; fátækur og umkomulaus, umkomulaus og einmana*). Vísbendingin um merkingarlega nálægð tiltekinna tveggja orða er fólgin í því hversu mörg orð þeim eru sameiginleg sem hinn hluti orðaparsins. Notkunardæmin eru að mestu sótt á vefsetrið *Tímarit.is* sem býr yfir afar ríkulegu efni til greiningar. Við samanburð á orðunum *glundroði* og *ringulreið* kemur t.d. í ljós að meðal förunauta þessara orða í orðapörum orðanestins (196 með *glundroði* og 217 með *ringulreið*) eru 102 orð sameiginleg báðum (t.d. bæði *glundroði* og *fálm*

og *ringulreið* og *fálm*). Að þessu leyti eru orðin tengdari innbyrðis en þau eru hvort um sig við önnur orð. Þegar tengslatíðnin er rakin áfram á bilinu 35–90 sameiginlegir förunautar er niðurstaðan fyrir *glundroði* þessi: upplausn 88, stjórnleysi 86, óstjórn 71, öngþveiti 62, ofbeldi 57, spilling 39, agaleysi 37. Orðið *ringulreið* tengist að mestu leyti sömu orðum á þessu bili: stjórnleysi 84, upplausn 84, öngþveiti 68, óstjórn 67, ofbeldi 65, vonleysi 43, agaleysi 38, hik 37, blóðsúthellingar 35.

Um leið og samanburðurinn staðfestir náinn merkingarskyldleika orðanna *glundroði* og *ringulreið* sýnir hann hvaða orð standa þeim næst hvoru um sig. Sé hins vegar litið til orðalistans í heild sem förunautarnir mynda og slíkum listum jafnvel slegið saman þegar merkingarskyld orð eiga í hlut myndast tengingar sem að meira eða minna leyti spanna heil merkingarsvið. Í því safni geta gægst fram orð sem við nánari athugun reynist full þörf á sem jafnheitum þótt þau leiti ekki á við fyrstu umhugsun.

7 Hugtaksleg nálgun og gagnkvæmni í lýsingu

Í texta tvímála orðabókar koma tengsl málanna fram í orðbundnu samhengi, þar sem viðfangsmálið myndar umgjörðina og ræður ferðinni en markmálið fylgir eftir með því að tefla fram samsvörunum við orð og orðasambönd viðfangsmálsins. Í þessari síendurteknu glímu við einstök orð eru hugtaka- og merkingarvensl málanna sífellt til skoðunar án þess að sú skoðun skilji eftir sig mikil ummerki í orðabókartextanum. Það skýrist að miklu leyti af því ójafnvægi sem er með málunum og að sú hugtaka- og merkingarflokkun sem á sér stað í orðabókarferlinu snertir aðeins annað málið hverju sinni og nær ekki að mynda heildir sem taka til beggja málanna.

Hér er á ferðinni einn helsti veikleiki tvímála orðabóka sem takmarkar notagildi þeirra í ýmsu samhengi (sbr. Anna Helga Hannesdóttir og Jón Hilmar Jónsson 2001 og Hanks 2012:60–61). Áherslan á jafngildi málanna og bein jafnheiti því til staðfestingar veldur því að fjölbreytni og breytileiki í mál- og orðanotkun skilar sér ekki sem skyldi, og orð og aðrar orðasafnseiningar sem ekki falla með góðu móti að jafnheitavenslum standa höllum fæti og verða hæglega út undan. Þetta á eðlilega í ríkari mæli við markmálið en fjölbreytileika viðfangsmálsins eru einnig takmörk sett.

7.1 Andspæni málanna í stafrænu umhverfi

Í stafrænu umhverfi eru aðstæður til að láta hið merkingar- og hugtakslega samhengi njóta sín betur og byggja upp lýsingu þar sem málin eru að meira eða minna leyti í jafnvægi. Snertipunktur málanna eru þá ekki jafngildin ein í formi orða heldur einnig og ekki síður hugtakavensl af ólíku tagi sem veita orðum og orðafari aðgang óháð því hvernig háttar til um bein jafnheiti. Slík lýsing getur að einhverju leyti sýnt að einstök hugtakasvið búa að misríkulegum orðaforða í ólíkum málum. Setningarleg og orðmyndunarleg sérstaða hvors máls um sig getur einnig komið mun skýrar fram.

Hugtaksbundin tvímála orðabókarlýsing af þessu tagi helst í hendur við þá merkingarflokkun og merkingargreiningu í hvoru máli fyrir sig sem lögð hefur verið áhersla á hér að framan sem undirstöðu nýrrar tvímála orðabókar. Hún leysir ekki orðbundna lýsingu af hólmi enda er notkunarskírskotunin að nokkru leyti ólík en með henni má veita skýrari innsýn í orðafar og orðanotkun og styrkja málbeitingarhlutverk orðabókarinnar.

7.2 Tvímála samanburður í ólíkum myndum

Víðtæk og rækileg forgreining á íslenskum orðabókargögnum skapar forsendur til að bera íslensku að öðrum tungumálum frá þessu sjónarhorni. Slíkur samanburður hefur sjálfstætt orðfræðilegt og orðabókarfræðilegt gildi en getur um leið verið liður í orðabókarverkefnum sem jafnframt er ætlunin að skili sér í hefðbundnari búningi. Í orðabókunum *Orðaheimi* (2002) og síðar *Stóru orðabókinni um íslenska málnotkun* (2005) er lagt inn á þessa braut gagnvart hugtakaflokkun íslenskra orðasambanda, þar sem íslensku hugtakaheitin í flettulista orðabókanna eiga sér samsvarandi heiti á ensku, dönsku og þýsku (í *Orðaheimi* er aðeins um að ræða ensku) sem fram koma í stafrófsröðuðum lista fyrir hvert tungumál. Með þessu er hugsað til þarfa erlendra notenda, m.a. nemenda í íslensku, sem vilja ná tökum á íslenskri málbeitingu og orðanotkun en hafa ekki sömu leitarforsendur og íslenskir notendur sem betur geta ratað að hugtakinu út frá tilteknu orðasambandi.

Hér beinist orðabókarlýsingin að einu máli og því ekki um markmál að ræða en erlendum notendum er opnuð leið að orðabókartextanum með eins konar markmálslegri skírskotun. Slíkar tengingar má mynda við fleiri mál og þær geta eins átt við um almennari hugtakaorðabók

sem nær í ríkari mæli til stakra orða. Stærra skref er að byggja upp tvímála lýsingu á grundvelli íslenska efnisins þar sem samstæður hugtakalisti tengir málin saman. Unnið hefur verið að íslensk-sænskri lýsingu af því tagi sem að vísu er enn ófullgerð en hefur sýnt fram á gildi þess að nálgast orðabókarlegan samanburð tveggja mála frá hugtakslegu sjónarhorni (sjá Anna Helga Hannesdóttir o.fl. 2010). Ætla má að óháð hugtakaskipan, þar sem jafnt er skírskotað til málanna tveggja, myndi enn betur tryggja jafnvægi í orðabókarlýsingunni.

8 Niðurlag

Hér hefur verið fjallað um efni og aðföng tvímála íslenskra orðabóka og hvernig aðstæður við gerð þeirra hafa mótað efnisskipan þeirra og framsetningu. Skortur á efnisgreiningu og efnisflokkun hefur komið í veg fyrir að nýjar orðabækur njóti í nægilegum mæli góðs af fyrri verkum og sá vandi blasir einkum við þegar horft er til eldri orðabóka í prentaðri mynd. Gögn sem liggja að baki endanlegum orðabókar-texta eru lítt eða ekki aðgengileg, og reynsla og innsæi orðabókarhöfunda skilar sér ekki á frjóan hátt í nýjum verkum.

Með tilkomu stafrænna gagna og veforðabóka hafa aðstæður breyst til hins betra að þessu leyti. Hægara er að halda utan um gögn sem gerð orðabókar byggist á og margvíslegt orðabókarefni er betur innan seilingar en áður. En enn meira máli skiptir að hægt er að greina og flokka orðabókarefni óháð tilteknum orðabókarnotum og án þeirrar ströngu takmörkunar sem jafnan gildir um orðabækur í endanlegri gerð og nýta slíka forgreiningu við vinnslu einstakra orðabóka.

Þörfin fyrir almenna greiningu og flokkun lýtur sérstaklega að merkingarþættinum og það á í ríkum mæli við um tvímála orðabækur hvort sem litið er til viðfangsmáls eða markmáls. Skýr yfirsýn um orð og orðafar á einstökum merkingarsviðum stuðlar að markvissum efnistöfum og áfangaskiptingu og þegar markmál er annars vegar eru merkingarvensl oft ein til leiðsagnar um viðeigandi jafnheiti.

Gagnvart notkunarumhverfi orða, lýsingu orðasambanda og vali og framsetningu notkunardæma er forgreining með áherslu á sjálfstæði fleirytra merkingarbærra eininga og samræmda framsetningu orðasambanda mikilvægur leiðarvísir. Í *Íslensku orðaneti* hefur farið fram umfangsmikil greining af því tagi sem tvímála íslenskar orðabækur geta notið góðs af. Samræmd framsetning er þar forsenda

fyrir árangursríkri röðun og flokkun efnisins burtséð frá því hvaða tilhögun gildir í orðabókarverkum sem kjósa að nýta sér orðanetið.

Veforðabækur búa yfir möguleikum til að jafna stöðu viðfangsmáls og markmáls í tvímála orðabókum, losa þær að nokkru undan hinum orðbundnu venslum þar sem annað málið er í forgrunni og byggja þess í stað upp jafnvægari lýsingu þar sem málin tengjast fyrst og fremst á hugtakslegum grunni. Slík lýsing leysir ekki hefðbundnari tvímála orðabækur af hólmi en hún mætir mikilvægum þörfum margra notenda, sérstaklega með tilliti til máltileinkunar og yfirsýnar um einstök svið orðaforðans.

Ritaskrá

- Anna Helga Hannesdóttir og Jón Hilmar Jónsson. 2001. Að hafa í sig og á. Ísländsk fraseologi i ett isländskt-svenskt perspektiv. *LexicoNordica* 8:67–91.
- Anna Helga Hannesdóttir, Sofia Tingsell og Jón Hilmar Jónsson. 2010. Mot en begreppsaserat isländsk och svensk fraseologisk ordbok. Reflektioner kring pragmatiska idiom. Harry Lönnroth & Kristina Nikula (ritstj.), *Nordiska studier i lexikografi 10. Rapport från Konferensen om lexikografi i Norden, Tammerfors 3–5 juni 2009*, bls. 140–149. Tammerfors.
- Anna B. Nikulásdóttir. 2012. Tölvutækur merkingarbrunnur fyrir íslenska máltækni. *Orð og tunga* 14:19–28.
- Det danske sprog- og litteraturselskab*. Vefsíður um tungumál: www.dsl.dk/sprog.
- Dönsk-íslensk orðabók*. 1992. Ritstjórar: Hrefna Arnalds og Ingibjörg Johannesen. Reykjavík: Ísafoldarprentsmiðja. (Einnig á vefnum: http://snara.is/bls/um/_dais.aspx.)
- Ensk-íslensk orðabók með alfræðilegu ívafi*. 1984. Reykjavík: Örn og Örlygur. (Einnig á vefnum: http://snara.is/bls/um/_enisal.aspx.)
- Ensk-íslensk skólaorðabók*. 1986. Jón Skaptason, ritstjóri. Reykjavík: Örn og Örlygur.
- Ensk-íslenska orðabókin*. 2006. Jón Skaptason, ritstjóri. Reykjavík: JPV útgáfa. (Einnig á vefnum: http://snara.is/bls/um/_enisjpv.aspx.)
- Frönsk-íslensk orðabók*. 1995. Þór Stefánsson, ritstjóri. Reykjavík: Örn og Örlygur. (Einnig á vefnum: http://snara.is/bls/um/_fris.aspx.)
- Hanks, Patrick. 2012. Corpus evidence and electronic lexicography. Í: Sylviane Granger & Magali Paquot (ritstj.), *Electronic Lexicography*, bls. 57–82. Oxford: Oxford University Press.
- Helgi Haraldsson. 1996. *Rússnesk-íslensk orðabók*. Reykjavík: Nesútgáfan.
- ISLEX-orðabókin. (e.d.) Þórdís Úlfarsdóttir, aðalritstjóri. Stofnun Árna Magnússonar í íslenskum fræðum. Á vefsíðunni www.islex.is. (Sótt 11. febrúar 2013.)

- Íslensk samheitaorðabók*. 2012. 3. útgáfa. Svavar Sigmundsson, ritstjóri. Styrktarsjóður Þórbergs Þórðarsonar og Margrétar Jónsdóttur. Reykjavík: Forlagið. (Einnig á vefnum: http://snara.is/bls/um/_samh.aspx.)
- Íslensk-ensk orðabók*. 1989. Sverrir Hólmarsson, Cristopher Sanders og John Tucker, ritstjórar. Reykjavík, Iðunn. [Önnur útgáfa, aukin og endurskoðuð. 2009. Christopher Sanders, ritstjóri. Reykjavík: Forlagið.] (Einnig á vefnum: http://snara.is/bls/um/_isen.aspx.)
- Íslensk-ítölsk orðabók*. 1994. Paolo Maria Turchi, ritstjóri. Reykjavík: Iðunn. (Einnig á vefnum: http://snara.is/bls/um/_isit.aspx.)
- Íslensk-spænsk orðabók*. 2011. Guðrún Tulinius o.fl., ritstjórar. Reykjavík: Forlagið. (Einnig á vefnum: http://snara.is/bls/um/_issp.aspx.)
- Íslensk-þýsk orðabók*. 1993. Björn Ellertsson, ritstjóri. Reykjavík: Iðunn.
- Íslensk-frönsk orðabók*. 1950. Gerard Boots, ritstjóri. Reykjavík: Ísafoldarprentsmiðja.
- Íslensk-rússnesk orðabók*. 1962. Valerij P. Bérkov, ritstjóri. Moskva.
- Íslenskt orðanet*. (e.d.) Jón Hilmar Jónsson og Þórdís Úlfarsdóttir. Stofnun Árna Magnússonar í íslenskum fræðum. Á vefsíðunni <http://ordanet.is/>. (Sótt 28. febrúar 2013.)
- Íslenskt textasafn*. Stofnun Árna Magnússonar í íslenskum fræðum. Á vefsíðunni <http://www.lexis.hi.is/corpus/leit.pl>.
- Íslenskur orðasjóður 2012*. Á vefsíðunni: http://wortschatz.uni-leipzig.de/ws_isl/.
- Ítölsk-íslensk orðabók*. 1999. Paolo Maria Turchi, ritstjóri. Reykjavík: Iðunn. (Einnig á vefnum: http://snara.is/bls/um/_itis.aspx.)
- Jón Hilmar Jónsson. 1995. Verbgrammatikk i islandske tospråklige ordbøker med islandsk som kildepråk. *LexicoNordica* 2:65–78.
- Jón Hilmar Jónsson. 2005. Aðgangur og efnisskipan í íslensk-erlendum orðabókum – vandi og valkostir. *Orð og tunga* 7:21–40.
- Jón Hilmar Jónsson. 2012. Að fanga orðaforðann: orðanet í þágu orðabókar. *Orð og tunga* 14:39–65.
- Mörkuð íslensk málheild*. Stofnun Árna Magnússonar í íslenskum fræðum: http://www.arnastofnun.is/page/morkud_islensk_malheild.
- Norsk-islandsk ordbok*. 1987. Hróbjartur Einarsson, ritstjóri. Oslo: Universitetsforlaget.
- Orðaheimur*. 2002. Jón Hilmar Jónsson, ritstjóri. Reykjavík: JPV útgáfa.
- Sanders, Cristopher. 2005. Bilingual Dictionaries of Icelandic: Types of Users and their Different Needs – a Discussion. *Orð og tunga* 7: 41–57.
- Snara vefbókasafn*. Á vefsíðunni: <http://snara.is/>.
- Spænsk-íslensk orðabók*. 2007. Guðrún Tulinius o.fl., ritstjórar. Reykjavík: Forlagið. (Einnig á vefnum: http://snara.is/bls/um/_spis.aspx.)
- Stóra orðabókin um íslenska málnotkun*. 2005. Jón Hilmar Jónsson, ritstjóri. Reykjavík: JPV útgáfa. (Einnig á vefnum: http://snara.is/bls/um/_stor.aspx.)
- Svensén, Bo. 2004. *Handbok i lexikografi*. Stockholm: Norstedts.

Svensk-isländsk ordbok. 1982. Gösta Holm og Aðalsteinn Davíðsson, ritstjórar. Lund: Bröderna Ekstrands Tryckeri AB.

Tímarit.is. Landsbókasafn Íslands – Háskólabókasafn. Á vefsíðunni: timarit.is. Þórdís Úlfarsdóttir. 2013. ISLEX – norræn margmála orðabók. Orð og tunga 15. 41–71. (Þetta hefti.)

Þýsk-íslensk orðabók. 2008. Heimir Steinarsson, ritstjóri. Reykjavík: Opna. (Einnig á vefnum: http://snara.is/bls/um/_deis.aspx.)

Þýsk-íslensk/íslensk-þýsk orðabók. 2004. Steinar Matthíasson, ritstjóri. Reykjavík: Iðnú.

Lykilorð

tvímála orðabækur, jafnheiti, merkingarleg vensl, hugtakaorðabók

Keywords

bilingual dictionaries, equivalents, semantic relations, thesaurus

Abstract

The paper describes the collection and presentation of Icelandic lexicographic material in bilingual Icelandic dictionaries, especially where Icelandic plays the role of the source language. The composition of the Icelandic lemma list is discussed where the lack of semantic analysis and classification has made it difficult to detect internal relations and to concentrate on words and phrases belonging to a specific concept or semantic field. Here the lexicographic description of Icelandic can benefit from a comprehensive analysis and presentation of the structural and semantic relations of the Icelandic vocabulary as found in the project *Íslenskt orðanet*. On the basis of a large collection of collocations the patterns of use reveal the interaction of semantic and syntactic relations, building groups of semantically coherent words and phrases. In accordance with the condition of monosemic lemmas the lemma list of *Íslenskt orðanet* contains many multi-word units, showing structural behavior of the lexical elements involved.

Sorting and classification of such lexicographic material which has gone through a proactive analysis, together with corpus evidence, will improve the usefulness of new bilingual Icelandic dictionaries in the future.

Jón Hilmar Jónsson
Stofnun Árna Magnússonar í íslenskum fræðum / Háskóli Íslands
Neshaga 16
107 Reykjavík
jhj@hi.is

Kristín Bjarnadóttir

Hvert á að sækja orðaforðann í orðabók?

1 Inngangur

Einn af grundvallarþáttum við gerð nýrrar orðabókar er að velja og takmarka flettiorðaforðann og vegna þessa má halda því fram að flettulistinn sé kjölfestan í hefðbundinni orðabók. Vegna tæknibreytinga hillir nú undir nýjar aðferðir til að nálgast orðaforðann og bylting hefur þegar orðið í aðgengi að orðmyndum í textum og greiningu þeirra. Í þessari grein er fjallað um mismunandi aðferðir við söfnun efnis í flettulistann. Þá er átt við orðalistann einan, án tillits til merkingargreiningar eða samhengis orða, og án tillits til útgáfuforms orðabókarinnar. Hér er eingöngu fjallað um aðferðir til þess að finna sjálft hráefnið, orð og orðmyndir, í þeim gögnum sem nú eru tiltæk.

Til skamms tíma var nánast aðeins ein leið fær við að afmarka orðaforðann í orðabók um íslenskt nútímamál: Að nýta lista sem fyrirrennararnir höfðu sett saman. Af útgáfusögu *Íslenskrar orðabókar* (Árni Böðvarsson 1963, 1983, Mörður Árnason 2000 o.áfr.; hér eftir Ío) er ljóst að þar bar hæst Blöndalsorðabók (sjá Guðrún Kvaran 1997) þar til Ío kom út en eftir það hefur síðara verkið verið meginstoð í íslenskri orðabókargerð. Í lýsingu Marðar Árnasonar ritstjóra 3. útg. Ío (2000) á fyrirhuguðum endurbótum á verkinu kemur skýrt fram það hlutverk orðabókar að draga saman efni úr öðrum áreiðanlegum heimildum og þar er fyrst að telja orðabækur:

... en nú sem fyrr hlýtur slík almenn íslensk orðabók að gegna því hlutverki að draga á handhægan hátt saman fróðleik úr sérhæfðum

Orð og tunga 15 (2013), 23–39. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

orðabókum, orðasöfnum, handbókum og fræðigögnum og vera sem almennast hjálpargagn fyrir sem allra flesta þátttakendur í íslensku málsamfélagi, vera „til gagns flestum þeim, sem þurfa skýringa á íslenskum orðum í almennu lesefni, gömlu eða nýju“ eins og það er orðað í formála fyrstu útgáfu (bls. V).

(Mörður Árnason. 1998:3)

Hér er ekki minnst á að orðabók ætti að gefa hugmynd um daglegt mál á vinnslutíma orðabókarinnar enda voru tæknilegar forsendur til þess ekki fyrir hendi hér á landi árið 1998 þegar Mörður skrifaði grein sína. Þar var enn byggt á þeim grunni sem notaður var við fyrstu útgáfu Ío, þ.e. Blöndalorsorðabók og söfn Orðabókar Háskólans (OH) sem sett var á stofn tveimur áratugum fyrr.

Orðabókarmönnum sem vinna að útgáfu orðabóka gefst yfirleitt takmarkað ráðrúm til að leggjast í orðtöku; það er einfaldlega of dýrt þegar horfst er í augu við tiltekinn útgáfudag sem venjulega er alltof skammt undan. Gildir þá einu hvort orðtakan er hefðbundin, þ.e. byggist á lestri texta, eða vélræn, þ.e. byggist á söfnun orðmynda úr rafrænum (en ómörkuðum) textum; báðar aðferðirnar eru seinlegar og matið á því hvað á heima í tiltekinni orðabók er erfitt.

Nýjar aðferðir í orðabókargerð byggjast á málheildum, þ.e. völdum textasöfnum sem ætlað er að gefa mynd af málinu eins og það er á hverjum tíma. Þar er m.a. stuðst við tíðnitölur til að meta hvort orð eiga heima í tilteknu orðabókarverki. Fyrsta enska orðabókin sem studdist við tíðnikannanir og málheild, *Cobuild*, kom út árið 1987.

Þar var markmiðið að birta enska tungu í orðabók, eins og hún væri raunverulega notuð (sjá bókarkápu: „... based on a detailed analysis of how today's English is really used“) en ekki aðeins eins og orðabókarhefðin segði til um:

For the first time, a dictionary has been compiled by the thorough examination of a representative group of English texts, spoken and written, running to many millions of words.

(Sinclair1987:XV)

Á Íslandi hefur slíkt orðabókarverk ekki enn litið dagsins ljós en reyndar er óljóst að nokkur málheild nái því að vera **staðgeng** (e. *representative*) þannig að sönn mynd af tungumálinu birtist í henni en það er forsenda þess að orðabók með þessu háleita markmiði verði gerð (Aðalsteinn Eyþórsson 2005:14). Samt sem áður er **jafnvæg** málheild (e. *balanced corpus*) ný gerð af heimild fyrir orðabókarmenn þar sem komast má nær orðaforða og málnotkun en hingað til hefur verið

hægt. Með tilurð *Markaðrar íslenskrar málheildar* (hér eftir MÍM) opnast því nýir möguleikar í íslenskri orðabókagerð (Sigrún Helgadóttir o.fl. 2012). Sama á reyndar við um önnur rafræn textasöfn, t.d. *Íslenskt textasafn* Stofnunar Árna Magnússonar í íslenskum fræðum¹ og *Íslenskan orðasjóð*² hjá Háskólanum í Leipzig (Erla Hallsteinsdóttir 2007, Quasthoff o.fl. 2012), þó að slík söfn séu ekki eins aðgengileg til úrvinnslu og MÍM.

Í þessari grein er fjallað um þá möguleika sem felast í þessum nýju gagnauppsprettum og gamlar og nýjar aðferðir bornar saman. Niðurstaðan er að á þessu stigi verður að nota allt efni sem tiltækt er til þess að notendur eigi möguleika á að finna sem flest af því sem þeir leita að. Íslenskir orðabókarnotendur eru fremur íhaldssamir í orðabókarnotkun og leita einna helst að merkingu orða og viðeigandi notkun þeirra (Ásta Svavarsdóttir 2008), en orðabókarmenn vita að væntingar notenda til orðabóka eru langt umfram það sem nokkur leið er að verða við. Þar þarf ekki að nefna annað en þá kröfu margra notenda að öll orð sem eiga sér viðurkenndan og fastan sess í málinu finnist í **orðabókinni**, án þess að tekið sé fram til hvaða orðabókar er vísað. Orð sem ekki finnast í „orðabókinni“ eru þá talin óbrúklek og óviðurkennd. Þessari kröfu geta orðabókarmenn seint sinnt að fullu. Engin orðabók verður nokkurn tíma tæmandi heimild um málið og hver á að taka sér það alræðisvald að ákveða hvaða orð mega lifa í málinu og hver ekki? Sem betur fer hafa orðabækur heldur ekki slíkt heljarafli í reynd en verkefnið er samt sem áður að sýna orðaforðann í málinu sem best, eins og hann er á hverjum tíma, og þess vegna er vert að skoða ný gögn og nýjar aðferðir.

Efnisskipan er þessi: Í 2. kafla er fjallað um hefðbundið íslenskt orðabókarefni, útgefnar orðabækur og önnur orðfræðileg gögn til orðabókargerðar sem unnin hafa verið með hefðbundnum hætti, þ.e. með hefðbundinni orðtöku. Þar er líka sagt frá Norræna verkefninu svokallaða þar sem gerð var tilraun til að sameina sem mest af þessu efni í rafrænan íslenskan orðabókarstofn. Í 3. kafla er fjallað um nýjar gagnalindir til orðabókargerðar, þ.e. rafræna texta, markaða og ómarkaða. Í 4. kafla er fjallað stuttlega um vélræna orðtöku. Í loka-kaflanum er fjallað um samspil gamalla og nýrra aðferða en niðurstað-

1 Í *Íslensku textasafni* eru gamlir og nýir textar, u.þ.b. 67 milljónir lesmálsorða. Niðurstöður leitar birtast sem orðstöðulykill, sjá http://arnastofnun.is/page/arnastofnun_gagnasafn_textasafn.

2 Í *Íslenskum orðasjóði* er nú u.þ.b. hálfur milljarður lesmálsorða af vefnum; sjá http://wortschatz.uni-leipzig.de/ws_ice/.

an er sú að slíkt verklag sé árangursríkasta leiðin við þær aðstæður sem nú eru.

2 Hefðbundið íslenskt orðabókarefni

Hefðbundið efni í íslenskri orðabókargerð er það samsafn sem orðabókarmenn taka í arf frá forverum sínum. Þar er aðallega byggt á útgefnum orðabókum, seðlasöfnum OH og orðalistum af ýmsu tagi úr ýmsum verkefnum. Orðabækur eru dýr langtímaverkefni og þess vegna verða orðabókarmenn að nýta allt tiltækilegt efni á sem allra ódýrastan máta. Þess er ekki kostur í litlu málsamfélagi að frumvinna flettiorðalista í hverja orðabók og þess vegna gengur hver orðabókarmaður í smiðju hjá öðrum. Fljótlegast er að vinna úr tölvutækum orðalistum en þar með er ekki nema hálf sagan sögð. Flettugildi í orðabók er ekki bundið við uppflettiórðin ein (sbr. Jón Hilmar Jónsson 2013), en hér er umfjöllunarefnið þó aðeins flettiorðalistinn sjálfur og hvar og hvernig best eða auðveldast er að nálgast efni í hann. Það efni sem orðabókarmenn taka í arf frá fyrirrennrum sínum skiptist í meginatriðum í tvennt: Í útgefnar orðabækur og gagnasöfn, upprunalega seðlasöfn, sem ætluð eru til orðabókargerðar.

2.1 Útgefnar orðabækur

Eins og fyrr segir eru meginstoðir í íslenskri orðabókargerð *Íslensk-dönsk orðabók* Sigfúsar Blöndals (1920–1924) og *Íslensk orðabók* (Ío 1963, 1983 o.áfr.) og byggir síðara verkið að miklu leyti á því fyrra. Heimildir að Blöndalsorðabók eru m.a. fornmálsorðabækur, t.d. orðabók Richards Cleasbys og Guðbrands Vigfússonar (Guðbrandur Vigfússon 1874–1876), og orðabók Björns Halldórssonar (1814) en til hennar var safnað í lok 18. aldar. Orðabókarmenn leita síðan fanga eins víða og nokkur kostur er og hér skal nefna *Íslensk-enska orðabók* eftir Sverri Hólmarsson, Christopher Sanders og John Tucker sem dæmi, en í formála þar eru nefndar fjölmargar tvímálaorðabækur sem leit að var í um orðaforða (Sverrir Hólmarsson o.fl. 1989). Fyrsta heimild sem nefnd er um orðaforðann er samt sem áður Ío: „Orðaforða höfum við meðal annars heyjað okkur úr Orðabók Menningarsjóðs, 2. útg. 1983, . . .“ (s.st.).³ Aðgangur að efninu úr Ío gjörbreyttist eftir að

3 Að auki er nefnt í formálanum að orðaforðinn hafi verið borinn saman við tíðnilista

gagnagrunnurinn Lexa varð til hjá Máli og menningu (hér eftir MM) fyrir tölvuútgáfuna árið 2000 og hefur sá gagnagrunnur verið meginstoðin í orðaforðanum í þeim verkum sem síðan hafa verið unnin hjá MM og síðar Forlaginu/JPV.

Orðalistinn úr Ío er ekki að öllu leyti heppilegur í tvímálaorðabók, bæði vegna þess að orðaforðinn þar er alltof mikill fyrir slíkt verk og einnig vegna þess að kröfurnar í nýju verki eru ekki þær sömu og í verki sem er að stofni til frá 1963 og var frá upphafi ætlað að þjóna sem alhliða orðabók fyrir íslenska notendur. Þarna koma fram í hnotskurn þeir gallar sem eru við að nota orðalista úr eldri orðabókarverkum:

- Skilgreining á markhópi og yfirgrípi á ekki við í nýju verki. Þetta á t.d. við í tvímálaorðabókum þar sem markmálið setur viðfangsmálinu skorður, m.a. vegna þess að þýðing getur verið nauðsynleg markmálsins vegna, þó svo að ekki sé þörf á merkingarlýsingu í viðfangsmálinu.
- Hluta orðaforðans vantaði til skamms tíma í orðabækur. Þetta á sérstaklega við um orð úr daglegu máli sem sjaldnast komast á bækur, t.d. heimilisorð ýmiss konar.
- Tilhneiging er til að safna sjaldgæfum orðum, á kostnað þess sem algengt er. Eitt af einkennum Ritmálssafns OH er t.d. betri yfirsýn um sjaldgæf orð og sérkennileg en algengt mál. Þar eru t.d. fjögur dæmi um orðið *eiginmaður* en níu dæmi um orðið *ektamaður*.
- Orðaforðinn úreltist og því þarf bæði að bæta við og grísa. Ýmis orð um tölvur eru þar nærtæk dæmi, t.d. *gataspjald* og *nálaprentari* sem nú eru úrelt, og *blogg* og *spjaldtölva* sem eru á hvers manns vörum.
- Villur og draugorð geta átt sér langa ævi. Orðið *nösgæs* 'e.k. gæs eða ef til vill önd' (Ío 1963, 1983) lifði t.d. lengi vel góðu lífi í orðabókum sem heiti á önd (eða gæs), meira að segja með latneska heitinu *anser nasutum*, enda þótt orðið sé úr gátu í Heiðreks sögu og merki andardráttur (Aðalsteinn Eyþórsson 1998).

Öll þessi atriði verða til þess að gaumgæfa þarf flettulistann sérstaklega í hverju nýju orðabókarverki en jafnframt þurfa orðabókarmenn að nýta þær heimildir sem til eru. Þá er vert að gera sér grein fyrir kostum og göllum þessara heimilda.

á lokastigi, þ.e. 7000 algengustu orðin í lista frá Eiríki Rögnvaldssyni og Vilhjálmi Sigurjónssyni. Ekki kemur fram við hvaða texta var stuðst í þeim lista.

2.2 Gagnasöfn Orðabókar Háskólans

Seðlasöfn OH eru frumheimildir Ío, ásamt Blöndalsorðabók, og þangað sækja orðabókarmenn enn efni. Þetta á einkum við um Ritmálssafnið sem er að stórum hluta aðgengilegt á vefnum til skoðunar. Sökum stærðar og takmarkana í leit er það þó fremur erfitt viðureignar en flettiorð þar eru yfir 700 þúsund. Talmálssafnið er enn sem komið er ekki aðgengilegt til úrvinnslu nema innan stofnunar. Efni úr báðum þessum söfnum var notað við gerð Ío, sem og einnig við gerð *Íslenskrar orðsifjabókar* (Ásgeir Blöndal Magnússon 1989). Bæði þessi söfn voru gerð með hefðbundnum aðferðum, orðtöku og spurningum til almennings. Til Ritmálssafns var safnað allt frá stofnun OH árið 1944 og rétt fram yfir 1990 en efni í Talmálssafn var einkum safnað í tengslum við þáttinn Íslenskt mál sem fluttur var í Ríkisútvarpinu 1956–2005. Eftir það hefur athyglinni verið beint að söfnun á rafrænum textum í *Íslenskt textasafn* (sjá kafla 3.1).

Hefðbundin orðtaka úr ritmáli felst í lestri á textum en þar birtist aðeins hluti af daglegu máli. Við söfnun efnis til Talmálssafnsins var yfirleitt leitað að tilteknum orðum eða lýsingum á fyrirbærum þannig að orðaforðinn var bundinn við tilteknar aðstæður eða orð, eins og liggur í hlutarins eðli í útvarpsþætti. Í þessar heimildir hafa íslenskir orðabókarmenn sótt sér efni í flettulista um langa hríð en gallinn er sá að efnið er bæði of mikið, þar sem orðaforðinn er gríðarlegur og því seinlegur til úrvinnslu, og of lítið, þar sem orð úr daglegu máli nútímans er ekki endilega að finna þarna. Þá er sá galli á söfnunum báðum að þar er safnað efni sem vekur athygli orðtökumannsins fyrir einhverra hluta sakir. Jaðartilvik, t.d. óvenjuleg málnotkun eða sjaldgæf orð, fá þá e.t.v. of mikið vægi á kostnað þess sem er almennt og algengt, sbr. dæmin *ektamaður* og *eiginmaður* sem nefnd voru hér að framan.

Önnur orðfræðileg gagnasöfn koma minna við sögu við gerð almennra nútímamálsorðabóka, þótt þau séu frábærar heimildir í sögulega orðabók, eins og t.d. seðlasafn um orðabók Jóns Ólafssonar úr Grunnavík,⁴ sem er góð heimild um mál 18. aldar. Íðorðaforðann er síðan að finna í Orðabanka Íslenskrar málstöðvar⁵ og þangað leita orðabókarmenn eftir því sem þörf er á. Ekki verður frekar fjallað um

4 Orðabókin er í handritinu AM 433 fol. Hún er útgefin en seðlar í uppskrift Jakobs Benediktssonar eru varðveittir á orðfræðisviði Stofnunar Árna Magnússonar í íslenskum fræðum og til er aðgengileg tölvuskra yfir uppflettiorðin (sjá: http://www.arnastofnun.is/page/ordabokarhandrit_jons_olafssonar).

5 Sjá: <http://www.ordabanki.hi.is/wordbank/search>.

Þessi söfn hér enda er sjónum einkum beint að því hvar nálgast má almennan orðaforða. Í Ío 1963 birtist fyrsta tilraun til að vinna úr efni í seðlasöfnum OH en í Norræna verkefninu svokallaða var aftur gerð tilraun til þess að sameina orðaforða úr gögnum OH til nota í orðabók.

2.3 Norræna verkefnið

Á árunum 1994–1996 var unnið að gerð íslensks orðabókarstofns sem ætlaður var sem grunnur að tvímála orðabókum milli íslensku og annarra Norðurlandamála. Verkefnið hlaut vinnuheimild Norræna verkefnið (Kristín Bjarnadóttir 1998). Upprunalega var ætlunin að orðabókarstofninn tæki til almenns orðaforða og að flettiorð í honum yrðu u.þ.b. 50 þúsund, úr lifandi almennu nútímamáli. Efni úr *Orðastað* Jóns Hilmars Jónssonar var notað sem grunnur en frumheimild þar er Ritmálssafnið (Jón Hilmar Jónsson 1994, formáli, bls. xxx). Síðan var bætt við meira efni úr Ritmálssafni og öðrum söfnum OH eftir þörfum. Allur orðaforði úr *Íslenskri orðtíðnibók* (Jörgen Pind o.fl. 1991) var hafður með og loks var stuðst við prentaðar orðabækur, s.s. Ío (1983) sem var höfð til hliðsjónar. *Íslensk orðtíðnibók* er byggð á textasafni með 500 þúsund lesmálsorðum og þaðan koma u.þ.b. 30 þúsund uppflettiorð, en annars er efnið í Norræna verkefninu fengið með hefðbundinni aðferð, þ.e. hér er efni úr eldri orðabókum og söfnum endurnýtt en þar liggur að baki hefðbundin orðtaka.

Til að bæta úr eyðum í orðaforðanum í Norræna verkefninu var nokkuð af textum orðtekið sérstaklega. Þar á meðal voru bæklingar um heilbrigðismál, félagsmál, tryggingar, bankakerfið, sálfræði, lögfræði fyrir almennung, sjávarútveg og tölvur, og námsbækur, t.d. í örverufræði, matvælafræði, vélsníði, húsgagnasmíði, íþróttum o.fl. Einnig voru gerðar tilraunir með orðtöku úr rafrænum textum, aðallega úr *Morgunblaðinu*.

Sigurborg Hilmarsdóttir stjórnaði orðtökunni og við hana unnu t.d. fjórir starfsmenn sumarið 1995 og söfnuðu u.þ.b. 2.500 orðum. Morgunblaðið reyndist drýgst í leit að orðum úr almennu máli. Í töflu 1 eru nokkur dæmi þaðan.

áherslupenni	grasrótarhreyfing	ljósleiðari
barnaís	heimasíða	margmiðlun
bleksprautuprentari	húsbréfalán	myndlykill
buxnadragt	kapalkefni	raðgreiðsla
eldislax	krókaleyfi	ríkisvixill
fjárlagahalli	leikskólakennari	talhólf
flísefni	líkamsklukka	

Tafla 1. Dæmi um orðtöku úr Morgunblaðinu.

Orðtakan var mjög tímafrek og var unnin með hefðbundnum aðferðum, þ.e. textarnir voru lesnir og leitað að orðum sem þóttu hafa gildi í orðabók. Virkar samsetningar voru ekki teknar með. Afraksturinn varð nokkur þúsund orð, dýrmæt og dýr.

Alls eru 241.784 flettiorð í Norræna verkefninu og eru þau flokkuð eftir orðgerð, þ.e. í ósamsett, afleidd og samsett orð, og eftir flettugildi, þ.e. eftir því hversu mikið erindi orðin eru talin eiga í orðabók. Þar er gerður greinarmunur á virkum samsetningum og orðgervingum (lexíkalíseruðum orðum) og einnig tekið tillit til tíðni og notkunar orða, eftir því sem gerlegt var á þeim tíma þegar verkið var unnið en þá var ekki um greiðan aðgang að tíðnitölum að ræða.

Í Norræna verkefninu var reynt að sameina allar heimildir sem tiltækar voru á OH á þeim tíma sem verkið var unnið. Þetta var í þann mund aðgangur að tölvutækum textum var að opnast en engar aðferðir voru tiltækar til þess að nýta þá til orðtöku umfram það sem gert var í *Íslenskri orðtíðnibók*. Ekki fékkst styrkur til þess að halda vinnu við Norræna verkefnið áfram árið 1996 og það lagðist í dvala. Þar varð samt til mikið orðasafn sem að mestu var tekið saman með hefðbundnum aðferðum og það er þess vegna gott til samanburðar við efni sem aflað verður með nýrri aðferðum. Þórdís Úlfarsdóttir (2013) tekur fram að 86% af orðaforðanum í ISLEX sé úr Norræna verkefninu þannig að efnið nýttist ágætlega í því verki.

3 Nýjar heimildir og nýjar aðferðir við orðtöku

Með tilkomu tölvutækra texta breyttist heimur orðabókarmannsins. Í stað hefðbundinnar orðtöku þar sem textar eru lesnir á pappír, kemur vélræn orðtaka sem byggist á því að láta tölvu finna orð eða orðmyndir í rafrænum textum sem ekki eru í samanburðarefni. Milli-stigið er að orðtaka úr rafrænum texta á skjá, með því að lesa textann,

en á þeirri aðferð og hefðbundinni lestrarorðtöku er bitamunur en ekki fjár. Talsverður vinnusparnaður er samt sem áður fólgin í þessari aðferð þar sem ekki þarf að skrifa efnið upp eða slá það inn. Í eiginlegri vélrænni orðtöku er markmiðið að fá yfirlit um allar orðmyndir í textanum en orðtakan verður markvissari eftir því sem greiningin á textunum er betri þar sem tvíræðni orðmynda er gríðarleg í íslensku.⁶

Sá grundvallarmunur er á vélrænni orðtöku og öðrum orðtöku-aðferðum að í vélrænu orðtökunni er allur orðaforðinn í viðkomandi texta undir; þar er unnið með tæmandi lista af orðum og orðmyndum. Þar þarf því að grísa og velja efnið í hvert orðabókarverk fyrir sig, eftir þeim aðferðum sem henta hverju sinni. Í hefðbundinni orðtöku er orðunum hins vegar safnað eftir tilfinningu orðtökumannsins, venjulega án mikils samanburðar við það efni sem fyrir er. Þar takmarkast umfangið á textunum af því sem er í mannlegum mætti að lesa og þess vegna er að einhverju leyti tilviljanakennt hvaða orð eru orðtekin.

Rafrænir textar eru að sjálfsögðu einnig notaðir til dæmaleitar í orðabókargerð. Með því að ganga út frá fyrirframgefnum orðalista og nota textasöfn eða málheildir til dæmaleitar fæst betra yfirlit yfir raunverulega orðanotkun og setningarlegt umhverfi en áður hefur verið kostur á. Hins vegar skilar slík vinna ekki yfirliti yfir heildar-orðaforðann í textunum nema með notkun orðtökutóls þar sem allar orðmyndir eru skoðaðar. Þarna á milli er verulegt stökk í tæknilegum efnum þar sem dæmaleit er möguleg án þess að fyrir liggi málfræðileg greining á textunum en orðtaka byggist á því að textarnir séu greindir og lemmaðir (þ.e. að gefið sé uppflettiorð eða lemma fyrir hverja orðmynd).

Helstu gagnalindir orðabókarmannsins nú um stundir eru því rafrænir textar, hverju nafni sem þeir nefnast, og *Mörkuð íslensk málheild* (MÍM), þar sem málfræðileg greining fylgir hverri orðmynd (Sigrún Helgadóttir o.fl. 2012). *Beygingarlýsing íslensks nútímamáls* (BÍN) gegnir hlutverki samanburðarefnis í þessu tilliti en BÍN er safn 270 þúsund beygingardæma, u.þ.b. 5,8 milljónir beygingarmynda með greiningarstrengjum (mörkum) (Kristín Bjarnadóttir 2012). Efni úr almennu máli í BÍN er að mestu leyti úr orðabókum og söfnum OH, þ.m.t. Norræna verkefninu.

⁶ Í *Beygingarlýsingu íslensks nútímamáls* (BÍN) hafa aðeins 31,5% orðmynda bara einn greiningarstreng (Kristín Bjarnadóttir 2012). Sjá nánar um BÍN hér á eftir.

3.1 Rafrænir textar

Á OH var byrjað að safna rafrænum textum um eða upp úr 1980. Safnið heitir *Íslenskt textasafn* og í því eru nú u.þ.b. 67 milljón lesmálsorð. Þar eru gamlir og nýir textar af ýmsum toga sem skiptast á milli 30 efnisflokka. Safninu fylgir ekki orðalisti og engin leið er að komast að því hver orðaforðinn er nema með málfræðilegri greiningu. Orðabókarmaðurinn þarf því að vita fyrir fram að hverju hann er að leita og getur ekki að óbreyttu notað safnið til orðtöku. Það er hins vegar frábær uppspretta dæma, með þeim annmarka þó að flokkun dæmanna getur verið mjög seinleg, m.a. vegna tvíræðni orðmynda. Í algengum orðum geta dæmi líka orðið óviðráðanlega mörg og þá getur verið djúpt á því sem leitað er að og erfitt að ná yfirsýn yfir efnið.

Önnur heimild sem vænlegt er að leita dæma í er t.d. tímaritasafn Landsbókasafns – Háskólabókasafns, *Tímarit.is*. Þar eru ljóslesnir textar, ásamt myndum af ritunum. Kosturinn við tímaritasafnið er að aldursmerking skilar sér vel með dæmunum en helsti ókosturinn er sá að í ljóslestrinum er talsvert af villum. Þá er vert að nefna *Íslenskan orðasjóð*, sem áður var minnst á (sjá nmgr. 2), sem er risavaxið safn af íslenskum vefsíðum en talsvert er af villum í því safni líka. Þessar heimildir nýtast best til leitar að tiltölulega sjaldgæfum orðum sökum þess hve seinlegt er að leita í þeim.

3.2 MÍM og BÍN

MÍM er textasafn með málfræðilegri greiningu sem unnið hefur verið að á OH og Stofnun Árna Magnússonar í íslenskum fræðum frá árinu 2002. Textarnir eru allir frá 2000–2009 og þeir eru úr eins fjölbreyttu efni og nokkur kostur er. Í MÍM eru u.þ.b. 25 milljón lesmálsorð. Hverju orði fylgir mark, þ.e. greiningarstrengur fyrir orðflokk og beygingarlegar upplýsingar. Markamengið er stórt, yfir 600 mörk (Sigrún Helgadóttir 2012).

Með því að bera saman MÍM og BÍN ætti að vera hægt að finna ný orð í málinu eða a.m.k. þau orð sem ekki hafa ratað til orðabókarmanna. Jafnframt finnast glöppur í BÍN sem þarf að uppfæra jafnóðum. Mikilvægasti kosturinn við MÍM til orðabókargerðar er sá að þar gefst í fyrsta sinn kostur á skilvirkri vinnu við orðalista með tíðnitölum sem t.d. er hægt að nota til flokkunar á uppflöttiorðum í orðabók. Með þessu ætti að vera tryggt að algengustu orð vanti ekki,

eins og stundum hefur viljað brenna við þegar stuðst er við tilfinningu orðabókarmannsins.

Mörkun og lemmun á MÍM er ekki lokið og þar leynast enn villur, bæði í uppflettimyndum og orðfloknum. Vegna þess liggja tölur um lemmufjölda og tíðni ekki fyrir enn. Hér á eftir er sagt frá tilraun til að áætla lemmufjölda í hluta af MÍM en hún byggist á samanburði orðmynda úr MÍM og beygingarmynda úr BÍN. Samanburðurinn var gerður haustið 2011 og takmarkið með honum var að bæta orðaforðanum úr MÍM við BÍN. Samkvæmt lauslegri áætlun eru u.þ.b. 125 þúsund lemmur í MÍM sem ekki finnast í BÍN.⁷

Í þeim hluta MÍM sem borinn var saman við BÍN eru u.þ.b. 17,5 milljónir lesmálsaorða. Þar af voru einkvæmar eða ótvíræðar orðmyndir (með marki) tæplega 738 þúsund og voru þær bornar saman við 5,8 milljónir beygingarmynda úr BÍN. Niðurstaðan úr samanburðinum er í töflu 2.

Orðmyndir í MÍM (með marki)	737.856
Þar af í BÍN	425.238
Þar af ekki í BÍN	312.618

Tafla 2. Orðmyndaaföldi í MÍM og BÍN.

Orðmyndir úr MÍM sem finnast í BÍN eru taldar vera tækar íslenskar orðmyndir, þ.e. 425.238 orðmyndir af 737.856. Orðmyndir sem ekki voru í BÍN voru skoðaðar hver fyrir sig á skjá, ríflega 312 þúsund orðmyndir. Af þeim eru u.þ.b. 60% íslenskar orðmyndir en 40% er annað efni, t.d. erlend orð (25%), villur (6%), skammstafanir (1,7%) o.fl. Þetta aðskotaefni eru u.þ.b. 13% af einkvæmum orðmyndum í MÍM en ekki liggur fyrir hve stór hluti af lesmálsaorðunum þetta er. Allar orðmyndir sem hafa verið aðlagðar að íslensku málkerfi flokkast sem „íslenskar orðmyndir“, þ.m.t. allar beygðar slettur (*pródúsjónin*) og beygð erlend nöfn, t.d. grísku nöfnin *Parmenídes* og *Erastopenes* (ef. *Parmenídesar*, *Erastopenesar*).

Niðurstaða af samanburðinum á MÍM og BÍN er að í MÍM séu u.þ.b. 125 þúsund lemmur sem ekki er að finna í BÍN. Samanlagður lemmufjöldi í BÍN og MÍM er því sennilega yfir 400 þúsund. Til þess að nýta þetta efni til orðabókargerðar þarf að grisja það, t.d. með því að styðjast við tíðnitölur úr MÍM þar sem orðaforðinn í BÍN er fenginn úr orðalistum sem hvorki styðjast við tíðnikannanir né neins konar

⁷ Samanburðurinn var unninn í samvinnu við Jón Friðrik Daðason og Sigrúnu Helgadóttur og kann höfundur þeim bestu þakkir fyrir.

flokkun eftir flettugildi til orðabókarnota. Grisjunin er allt annað en einföld þar sem tíðnitölurnar einar duga ekki. Orðið *doppóttur* kemur t.d. aðeins þrisvar sinnum fyrir í MÍM og orðið *rykking* alls ekki en samt eru orðin síður en svo sjaldgæf í umræðu um saumaskap. Orðaforði af því sviði tilheyrir reyndar þeim hluta orðaforðans sem til skamms tíma vantaði í orðabækur, þ.e. heimilisorð ýmiss konar eða orð af verksviði kvenna. Orð sem koma sjaldan fyrir í rafrænum textum geta því vissulega átt heima í orðabók og orðin tvö sem hér eru nefnd eru bæði uppflettiorð í Ío og í ISLEX.

Þegar tíðni orðanna í MÍM sem ekki eru í BÍN er skoðuð koma strax í ljós gloppur í hefðbundna orðabókarefninu sem BÍN byggist að mestu leyti á. Dæmi um efni af þessu tagi eru í töflu 3 en þar eru m.a. algeng orð úr MÍM sem hljóta að eiga heima í orðabókum.⁸

	Orðmynd	Mark	Lemma	Ofl.
1262	vísindavefnum	nkeþgs	vísindavefur	kk
282	sonu	nkfo	sonur	kk
258	grunnskólabekk	nkeþ	grunnskólabekkur	kk
231	r-listans	nkeegs	r-listi	kk
177	knattspyrnustjóri	nken	knattspyrnustjóri	kk
158	skjöldu	nveo	skjöldur	kk
151	vísindavefsins	nheegs	vísindavefur	hk>kk
141	ameríkufari	nken-s	ameríkufari	kk
126	íbúðalánasjóðs	nkee-s	íbúðalánasjóður	kk
122	r-listinn	nken-s	r-listinn	kk.mgr.
111	vopnfirðinga	nkfe-s	vopnfirðingur	kk
101	ríkislögreglustjóra	nkee	ríkislögreglustjóri	kk
99	óskarinn	nkeng	óskar	kk
96	madur	nken	maður	kk
95	mogganum	nkeþgs	moggi	kk
93	danakonungur	nken-s	danakonungur	kk
93	danakonungs	nkee-s	danakonungur	kk

Tafla 3. Orðmyndir í MÍM sem ekki eru í BÍN.

Vandi orðabókarmannsins við úrvinnslu á efni úr MÍM er tvennis konar. Í fyrsta lagi er massinn gríðarlegur og seinlegt að vinna úr öllu efninu, jafnvel þótt öll greining væri rétt. Það erfiði er þó fremur fólgið í vinnu við merkingargreiningu en beinlínis í því sem hér er til

⁸ Í töflunni eru einnig úreltar beygingarmyndir sem ekki finnast í BÍN þó að þær komi fyrir í formlegu máli (*sonu*, *skjöldu*) og stafsetningarafbrigði sem finnast í tölvumáli þar sem séríslenskum stöfum er sleppt (*madur*). Í samkeyrslunni var ekki gerður greinarmunur á hástaf og lágstaf í samkeyrslunni (*r-listi*, *danakonungur*).

umræðu, þ.e. vinnu við orðalistann sjálfan. Í öðru lagi er orðaforðinn ekki eins mikill og æskilegt væri, eins og sjá má af orðinu *rykking* sem nefnt var hér að framan og á samanburði á orðamyndafjöldanum í MÍM og uppflettiorðafjölda í hefðbundnu orðabókarefni, t.d. Ritmálssafni, en MÍM skilar aðeins broti af þeim orðaforða sem þar er.⁹

Textasafn sem ætti að skila öllum þeim orðaforða sem venjulegir málnotendur hafa að einhverju leyti á valdi sínu þyrfti að vera gríðarlegt, án þess að því sé haldið fram að allur sá orðaforði sé á færi hvers manns. Þetta er einmitt efni sem ætla má að notendur vilji leita að í orðabók.

4 Vélræn orðtaka

Þær hugmyndir hafa verið uppi um allnokkurn tíma að gott væri að orðtaka efni jafnóðum og það birtist, t.d. á vefsíðum dagblaða. Með BÍN fylgir orðtökutól og með því er hægt að finna orðmyndir í textum sem ekki eru í BÍN. Tólið er enn á tilraunastigi og er frumstætt; það lemmar ekki og skilar býsna óaðgengilegum listum af orðmyndum (og öðrum strengjum). Úrvinnslan er því mjög seinleg. Til þess að tólið komi að fullu gagni þarf málfræðilega greiningu á textunum (mörkun), lemmun og síu fyrir rusl, samanber tölur um aðskotaefni í samanburðinum á orðaforðanum í MÍM og BÍN hér að framan. Með sjálfvirkri mötun á textum, t.d. dagblöðum hvers dags, væri þannig hugsanlegt að finna ný orð sem birtast frá degi til dags, eða a.m.k. orð sem ekki hafa ratað á fjörur orðabókarmanna fyrr.

Þegar orðtökutólið var prófað á stuttan texta um orðabókarfræðilegt efni var afraksturinn fjögur uppflettiorð (orðmyndir í sviga) og niðurstaðan er býsna dæmigerð, þ.e. fjögur samsett orð:

gagnaefni (gagnaefnis), *einmála* (einmála), *orðabókareining* (orðabókar-eininganna), *orðabókargögn* (orðabókargögn, orðabókargagna).

Íslensk orðmyndun er afskaplega virk og tilraunir með tólið sýna að samsetningar eru líklegasta afurðin í venjulegum texta. Vandí orðabókarmannsins er að finna hvaða samsetningar eiga erindi í orðabók, þ.e. þarna þarf að sigta úr þær samsetningar sem ekki eru fullkomlega gagnsæjar (Kristín Bjarnadóttir 1995).¹⁰

⁹ Ritmálssafnið er að vísu efni til sögulegrar orðabókar, allt aftur til 16. aldar.

¹⁰ Í tvímálaorðabók þarf einnig að taka tillit til þýðingarorðanna.

Virgni og einkenni íslenskrar orðmyndunar verða einnig til þess að tíðnitölur orða og sjálfstæðra orðmynda skila e.t.v. ekki eins afgerandi niðurstöðum fyrir íslensku og t.d. fyrir ensku, eins og sjá má af þessu dæmi:

kransæðarhjáveitugræðlingur
coronary artery bypass graft

Þarna er eitt orð í íslensku fyrir fjögur í ensku. Íslensku uppflettiorðin sem vert er að skoða eru *kransæð*, *hjáveita*, *græðlingur* og/eða *kransæðarhjáveitalhjáveitugræðlingur*. Vandi orðtökumannsins er því sá að einnig þarf að skoða orðhlutana til þess að grandgæfa orðaforðann í textunum, samanber þetta samsetta orð. Forrit sem greinir samsett orð þarf því að vera hluti af orðtökubúnaðinum¹¹ og tíðni orðhlutanna þarf sennilega einnig að koma við sögu ef nota á orðtíðni í íslenskri orðabókargerð. Það verk biður þó síðari tíma þegar orðtökutólið hefur batnað.

5 Lokaorð

Nýjar gagnauppsprettur gera orðabókarmönnum fært að endurskoða orðaforðann í verkum sínum á skilvirkari hátt en áður hefur verið mögulegt þannig að betur sé komið til móts við notendur með því að ná yfir fjölbreytta málnotkun á þeim tíma sem orðabókin er samin. Hins vegar eru ýmsir annmarkar á því að byggja íslenska orðabók alfarið á tíðnitölum, eins og gert var í *Cobuild*-orðabókinni sem nefnd er í upphafi þessarar greinar. *Cobuild* er byggð á málheild með u.þ.b. 26 milljónum lesmálsorða en lesmálsorð í MÍM eru u.þ.b. 25 milljónir í endanlegri gerð. Í *Cobuild* munu vera „over 70,000 references“ eða uppflettimyndir (sjá bókarkápu) en samkvæmt ágiskun eru u.þ.b. 300 þúsund uppflettiorð í MÍM. Ætla mætti að svo mörg uppflettiorð skilðu niðurstöðu sem hægt væri að byggja á til orðabókargerðar en mjög virkar samsetningarreglur verða til þess að tíðnitölur í ensku og íslensku eru ekki sambærilegar, eins og vikið var að hér að framan. Vegna þessa verður tíðnidreifing í íslensku ekki sambærileg við ensku. Til dæmis koma rúmlega 308 þúsund orðmyndir úr þeim hluta MÍM sem borinn var saman við BÍN aðeins fyrir einu sinni.¹² Stakdæmi

¹¹ Jón Friðrik Daðason (2012) notar forrit sem greinir samsett orð í leiðréttingarbúnaði sínum en tólinu hefur ekki verið beitt í orðtöku enn sem komið er.

¹² Fjöldi lemma í þessum hluta MÍM er 407.192 en af þeim koma 229.895 aðeins

eru veikur grunnur fyrir uppflöttiorð í orðabókum og m.a. þess vegna er nauðsynlegt að skoða tíðni orðhlutanna líka. Tilgáta mín er því að orðaforðinn úr MÍM myndi skila gloppóttri orðabók og að hefðbundinn efniviður íslenskra orðabóka, orðalistar og gagnasöfn, sé einnig nauðsynlegur.

Eftir þetta yfirlit um efni og aðferðir til að velja orðaforða í íslenska orðabók er niðurstaðan þessi:

- Hefðbundnir orðalistar eru enn sem komið er ómissandi. Þá þarf að grisja og uppfæra með nýjum aðferðum.
- Vélræn orðtaka er ekki enn komin í gagnid en mun nýtast til að hafa fingurinn á púlsinum og finna ný orð.
- MÍM og önnur stór rafræn söfn munu verða gríðarlega mikilvæg heimild um orðaforðann. Til þess að þau komi að fullum notum þarf að greina samsett orð í sundur.
- Engar vélrænar aðferðir eru enn til til þess að meta flettugildi orða. Þar þarf orðabókarmaðurinn að meta efnið á hefðbundinn seinlegan hátt, þótt hægt sé að styðjast við tíðnitölur, með þeim annmörkum sem á þeim eru í því efni sem nú er til reiðu. Textasöfnin eru einfaldlega ekki enn nægilega yfirgripsmikil til þess að innihalda allan þann orðaforða sem birtist í heimildum sem safnað er til með gamla laginu.

Heimildir

- Aðalsteinn Eyþórsson. 1998. Síðasti furðufuglinn — eða gæsin sem stóð á öndinni. *Frejas Psalter til brug for Jonna Louis-Jensen*, bls. 5–7. København. (Einnig birt á *Kistunni* 20. maí 2002: www.kistan.is; sótt 12.3.2013.)
- Aðalsteinn Eyþórsson. 2005. Hver er kjarni orðaforðans? *Orð og tunga* 7:9–20.
- Árni Böðvarsson (ritstj.). 1963. *Íslensk orðabók* handa skólum og almenningi. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Árni Böðvarsson (ritstj.). 1983. *Íslensk orðabók*. 2. útg. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Ásgeir Blöndal Magnússon. 1989. *Íslensk orðsifjabók*. Reykjavík: Orðabók Háskólans.
- Ásta Svavarsdóttir. 2008. Ordbogsbrug og ordbogskultur i Island. Under-søgelse of ordbogsbrug blandt islandske modernmåslærere. *LexicoNordica* 15:115–134.

einu sinni fyrir. Lemmunin er óleiðrétt, eins og áður kom fram, og tölurnar munu breytast í endanlegri gerð.

- Beygingarlýsing íslensks nútímamáls*. Ritstjóri Kristín Bjarnadóttir. <http://bin.arnastofnun.is>.
- Björn Halldórsson. 1814. *Lexicon Islandico-Latino-Danicum Björnionis Haldorsonii*, I–II. Havniæ.
- Blöndalsorðabók = Sigfús Blöndal. 1920–1924. *Íslensk-dönsk orðabók*.
- Cobuild = Sinclair, John (ritstj.). 1987. *Collins Cobuild English Language Dictionary*.
- Erla Hallsteinsdóttir. 2007. Íslenskur orðasjóður. *Orð og tunga* 7:109–124.
- Guðbrandur Vigfússon. 1874–1876. *An Icelandic-English Dictionary based on the MS. Collections of the late Richard Cleasby*. Oxford.
- Guðrún Kvaran. 1997. Rætur og heimildir. *Orð og tunga* 3:9–14.
- Ío = *Íslensk orðabók* (sjá Árne Böðvarsson 1963, 1983; Mördur Árnason 2000.)
- Íslenskur orðasjóður*. http://wortschatz.uni-leipzig.de/ws_ice/.
- Íslenskt textasafn*. http://arnastofnun.is/page/arnastofnun_gagnasafn_textasafn.
- Jón Friðrik Daðason. 2012. *Post-Correction of Icelandic OCR Text*. MS ritgerð í tölvunarfræði við Háskóla Íslands. (Sjá: <http://hdl.handle.net/1946/12085>.)
- Jón Hilmar Jónsson. 1994. *Orðastaður*. Reykjavík: Mál og menning.
- Jón Hilmar Jónsson. 2013. Að byggja undir íslensk-erlenda orðabók. Forgreining og orðabókarefni. *Orð og tunga* 15:1–22.
- Jörgen Pind, Friðrik Magnússon og Stefán Briem. 1991. *Íslensk orðtíðnibók*. Reykjavík: Orðabók Háskólans.
- Kristín Bjarnadóttir. 1995. Lexicalization and the Selection of Compounds for a Bilingual Icelandic Dictionary Base. Í: Ásta Svavarsdóttir, Guðrún Kvaran, Jón Hilmar Jónsson (ritstj.), *Nordiske studier i leksikografi* 3, bls. 255–263. Reykjavík: NFL, Nordisk språksekretariat og Orðabók Háskólans.
- Kristín Bjarnadóttir. 1998. *Norræna verkefnið*. Skýrsla um íslenskan orðabók-arstofn, 6. mars 1998. <http://www.lexis.hi.is/kristinb/norr.pdf>.
- Kristín Bjarnadóttir. 2012. The Database of Modern Icelandic Inflection. *Proceedings of "Language Technology for Normalization of Less-Resourced Languages"*, SaLTMiL 8 – AfLaT 2012. Istanbul, Tyrklandi.
- Mördur Árnason. 1998. Endurútgáfa *Íslenskrar orðabókar*: Stefna – staða – horfur. *Orð og tunga* 4:1–8.
- Mördur Árnason (ritstj.). 2000. *Íslensk orðabók*, tölvuútgáfa. Reykjavík: Edda hf.
- Orðabanki Íslenskrar málstöðvar*. Ritstjóri Ágústa Þorbergsdóttir. <http://www.ordabanki.hi.is/wordbank/>.
- Mörkuð íslensk málheild*. Ritstjóri Sigrún Helgadóttir. <http://mim.hi.is>.
- Ritmálssafn Orðabókar Háskólans*. http://www.arnastofnun.is/page/gagnasofn_ritmalssafn.
- Sigfús Blöndal. 1920–1924. *Íslensk-dönsk orðabók*. Reykjavík.
- Sigrún Helgadóttir, Ásta Svavarsdóttir, Eiríkur Rögnvaldsson, Kristín Bjarna-

- dóttir og Hrafn Loftsson. 2012. The Tagged Icelandic Corpus (MÍM). *Proceedings of "Language Technology for Normalization of Less-Resourced Languages"*, SaLTMiL 8 – AfLaT 2012. Istanbul, Tyrklandi.
- Sinclair, John (ritstj.). 1987. *Collins Cobuild English Language Dictionary*. London and Glasgow: Collins.
- Sverrir Hólmarsson, Christopher Sanders, John Tucker. 1989. *Íslensk-ensk orðabók*. Reykjavík: Iðunn.
- Tímarit.is*. Landsbókasafn Íslands – Háskólabókasafn: <http://timarit.is>.
- Quasthoff, Uwe, Sabine Fiedler, Erla Hallsteinsdóttir. 2012. *Frequency Dictionary. Icelandic. ISL*. Leipzig: Leipziger Universitätsverlag.
- Þórdís Úlfarsdóttir. 2013. ISLEX – Norræn margmála orðabók. *Orð og tunga* 15:41–71. (Þetta hefti.)

Lykilorð

orðabókarfræði, uppflöttiorð, orðaforði, málheild, vélræn orðtaka

Keywords

lexicography, headwords, vocabulary, corpora, automatic excerption

Abstract

The topic of this paper is different methods in assembling a list of headwords for Icelandic dictionaries. Until recently, the classic Icelandic dictionary (*Íslensk orðabók*) originally published in 1963, has been the primary source of data of the vocabulary for Icelandic lexicographers, with the collections at the department of lexicography of The Árni Magnússon Institute for Icelandic Studies as additional sources. With the advent of electronic text collections, notably the tagged Icelandic corpus (*Mörkuð íslensk málheild*, MÍM), Icelandic lexicographers now have access to a huge new source of data. The use of data on frequency from these new sources is of great value to lexicographers in the choice of headwords, but the complete coverage of the words and word forms in the texts made possible by new methods of word extraction also complements older material by filling accidental gaps in material assembled by older methods of dictionary excerption. Partly because of data scarcity problems in a language with a very rich morphology, the conclusion in the paper is that word frequency alone cannot be the base of an Icelandic dictionary at this point in time, as the volume of texts needed for a good coverage of the vocabulary far exceeds anything available now. One of the main reasons for this is a very productive system of compounding, making it necessary to split compounds before attempting to use frequency for the selection of vocabulary.

Kristín Bjarnadóttir
Stofnun Árna Magnússonar í íslenskum fræðum / Háskóli Íslands
kristinb@hi.is



Þórdís Úlfarsdóttir

ISLEX – norræn margmála orðabók

1 Inngangur

ISLEX er samnorrænt, rafrænt orðabókarverk sem er unnið hjá orðfræðisviði Stofnunar Árna Magnússonar í íslenskum fræðum (SÁ) (áður Orðabók Háskólans) í samstarfi við fjórar aðrar háskóla- og fræðastofnanir á Norðurlöndum, í Danmörku, Noregi, Svíþjóð og Færeyjum. Samstarfsaðilar verkefnisins eru *Det Danske Sprog- og Litteraturselskab* (DSL) í Kaupmannahöfn, *Institutt for lingvistiske, litterære og estetiske studier* við Háskólann í Bergen, *Institutionen för svenska språket* við háskólann í Gautaborg og *Fróðskaparsetur Føroya* í Þórs-
höfn.

ISLEX er veforðabók milli íslensku og fimm annarra norrænna mála. Hún hefur að geyma tæplega 50.000 íslensk uppflettiorð en markmálin eru danska, norska (bókmál og nýnorska), sænska og færeyska. ISLEX er eina norræna orðabókin hingað til sem sameinar svo mörg mál í einu verki. Hún er fyrst og fremst orðabók yfir nútímaíslensku og taka uppflettiorð, notkunardæmi og orðasambönd einkum mið af málnotkun samtímans. Verkinu er ætlað að svara brýnni þörf fyrir handhægar orðabækur milli íslensku og hinna norrænu málanna þar sem þær bækur sem til eru hafa flestar úrelst og uppfylla ekki lengur kröfur tímans. En um leið er markmið aðstandenda verkefnisins að búa til fjölþætta, nýstárlega og vandaða orðabók sem nýst getur sem flestum, Íslendingum og útlendingum, tungumálanemendum, kennurum og þýðendum, og raunar öllum þeim sem þörf hafa fyrir slíkt verk. Vinnan við verkið stóð að mestu leyti yfir á árunum 2006–2011 og var ISLEX

Orð og tunga 15 (2013), 41–71. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

opnað fyrir almennan aðgang í nóvember 2011, enda þótt ýmsum verkþáttum væri þá ekki fulllokið.

Í þessari grein er fjallað um ISLEX-verkefnið og er megináhersla lögð á þá þætti sem snúa að íslenska hluta orðabókarinnar. Efninu er skipað á eftirfarandi hátt: 2. kafli lýsir forsögu og upphafi verkefnisins, 3. kafli fjallar almennt um rafrænu orðabókina, 4. kafli greinir frá verklagi við ritstjórn í fjölþjóðlegu samstarfi, í 5. kafla segir frá efni og uppbyggingu orðabókarinnar, í 6. kafla er lýst notendaviðmóti hennar á vefnum og í 7. kafla er fjallað um ýmsa þá möguleika sem búa í svona verki og hugað að framhaldi þess. Lokaorð eru í 8. kafla.

2 Upphaf verkefnisins

Upphaf ISLEX-verkefnisins má rekja til tíunda áratugarins, um það bil áratug áður en byrjað var á sjálfri orðabókarvinnunni.¹ Árið 1994 réðst Orðabók Háskólans (OH) í samvinnu við Norræna málstöð í Osló í að semja íslenskan orðabókarstofn sem orðið gæti grunnur að tvímála orðabókum milli íslensku og annarra Norðurlandamála. Til verksins fékkst styrkur til tveggja ára frá Norræna menningarsjóðnum. Orðabókarstofninn skyldi vera um 50–70 þúsund orð. Stofninn var listi yfir valin orð í íslensku með áherslu á nútímamálið, en hverju orði fylgdu upplýsingar um orðflokk, formgerð (grunnorð eða samsett orð) og heimild um orðið. Kristín Bjarnadóttir hafði umsjón með þessu verki (sjá Kristín Bjarnadóttir 1998).

Lengi hafði verið vilji til þess að setja saman nútímalega íslensk-sænska orðabók, bæði á Íslandi og í Svíþjóð, og árið 2002 var leitað samstarfs við *Institutionen för svenska språket* við Gautaborgarháskóla um gerð slíkrar orðabókar. Áhugi var á því að koma á sams konar samvinnuverkefni við Norðmenn um íslensk-norska orðabók, og árið 2003 fékkst fjárstyrkur frá Det kongelige kultur- og kirke departement í Noregi til undirbúnings þess. Norski samstarfsaðilinn í verkefninu varð *Nordisk institutt* við háskólann í Bergen. Norðmenn lögðu áherslu á að bæði norskt bókmál og nýnorska skyldu vera með í orðabók af þessu tagi en ekki bókmálið eingöngu. Frá upphafi var jafnframt ljóst að það yrði á allan hátt hentugt að bæði tungumálin, sænska og norska, yrðu unnin samhliða í sama tölvukerfinu.

Haldin var frumkynning á orðabókinni í Kaupmannahöfn 2005, og

1 Sjá nánar um forsögu ISLEX: Þórdís Úlfarsdóttir 2011, Svavar Gestsson 2011 og Anders Ljuggren 2011.

í kjölfarið lýstu fulltrúar frá *Det Danske Sprog- og Litteraturselskab* (DSL) yfir vilja sínum til þess að taka þátt í ISLEX með dönsku sem markmál. Eftir frekari viðræður varð það úr að DSL sótti um fjárveitingu til að geta hafið verkið, og fékkst hún. Þar með voru fjögur lönd orðin aðilar að verkefninu.

Þátttaka Færeyja kom til sögunnar í ársbyrjun 2011 en þá var verkið komið langt á veg í hinum löndunum.

Hvert land hefur séð um kostnaðarhliðina við sitt tungumál. Verkið hefur að mestu verið fjármagnað með opinberu fé í öllum þátttöku-löndunum, auk styrkja úr samnorrænum og öðrum sjóðum.

3 Rafræn orðabók

Þegar byrjað var á verkefninu höfðu runnið upp nýir tímar í íslenskri orðabókargerð, tölvutæknin var stöðugt í þróun og ekki síst hafði veraldarvefurinn og allt umhverfi hans tekið stórfelldum framförum á liðnum áratug. Á fáum árum höfðu einnig kröfur umhverfisins aukist um að sem mest af upplýsingum væru aðgengilegar á vefnum, gjarnan ókeypis. Það er ekki að öllu leyti einfalt í orðabókarvinnu að svara þessum kröfum vegna þess að hvort sem birtingarformið er rafrænt eða ekki er starfið í eðli sínu afar seinlegt, og hætt er við að tæknileg framsetning sem hönnuð er utan um orðabókarverk snemma á ferli þess verði óþægilega fljótt úrelt eða gamaldags. Reynt hefur verið að mæta þessum vanda í ISLEX með því að endurbæta og byggja utan á gagnagrunn og notendaviðmót orðabókarinnar eftir því sem aðstæður og fjárhagur hafa leyft.

ISLEX-orðabókin er aðeins til sem rafrænt verk en hún var frá upphafi ætluð til birtingar á vefnum. Að sumu leyti gilda önnur lögmál um rafræna orðabók en prentaða bók. Rafrænt verk býður upp á meiri hraða í upplýsingaöflun, meira myndefni, hreyfimyndir og hljóð, tengla innan verksins og út fyrir það (í aðrar vefsíður). Einnig hefur rafræn bók þann kost að stöðugt er hægt að gera endurbætur og breytingar sem skila sér strax til notandans. Rafræn bók hefur enn fremur möguleika á miklu meiri útbreiðslu en prentuð bók. En fyrir utan sjálfa orðabókina hefur starfið við ISLEX skilað dýrmætu gagnasafni með skipulegum upplýsingum um vensl norrænna tungumála sem nýta má í öðrum tilgangi, bæði hagnýtum og fræðilegum.

ISLEX-verkefnið er unnið í sérsmíðuðum gagnagrunni með viðmót fyrir vefbirtingu og er tölvukerfið rekið hjá SÁ. Sérstakt viðmót er

fyrir ritstjórnarvinnu og annað sem er ætlað til uppflettinga, þ.e. notendaásýnd orðabókarinnar. Þessi tvö viðmót eru gjörólík en bæði sækja efnið í einn og sama gagnagrunninn.

4 Verklag við ritstjórn

Fjórar til fimm ritstjórnir í jafnmörgum löndum hafa að jafnaði starfað við ISLEX-verkefnið. Hver hópur er sjálfstæð starfseining sem hefur séð um sitt tungumál í orðabókinni. Eðli málsins samkvæmt er starf íslensku ritstjórnarinnar allólíkt starfi hópanna í Skandinavíu en þeirra verkefni er fyrst og fremst að þýða íslenskt efni orðabókarinnar á hin málin. Íslenski hópurinn hefur séð um að gera viðfangsmálinu skil og búa það í hendur ritstjóra markmálanna. Vinnan við verkefnið hefur ætíð farið fram samtímis í öllum löndunum, með þeirri undantekningu að Færeyingar bættust ekki í hópinn fyrr en árið 2011 eins og áður segir.

Starfandi er samráðsnefnd allra þátttökulanda sem hefur það hlutverk að annast formleg tengsl milli samstarfsstofnananna og eru meðlimir hennar jafnframt opinberir fulltrúar verkefnisins fyrir hönd sinna stofnana.

4.1 Verklag á Íslandi

Íslenska er viðfangsmál ISLEX-orðabókarinnar og er úrvinnsla þess margslungið starf en í grófum dráttum felst hún einkum í eftirfarandi þáttum. Orðum er skipt upp í merkingarliði þar sem það á við, samin er orðskýring á íslensku við hvern lið, notkunardæmi eru búin til eftir þörfum, orðasambönd eru færð í staðlað form og samin við þau skýring. Sérmerkingum og millivísunum er komið fyrir þar sem þurfa þykir, og sums staðar er mynd bætt inn í flettugreinina. Ástæður þess að íslenskum skýringum er komið fyrir í flettugreinunum eru ýmsar, ein er sú að með því fæst skýrari sundurgreining á merkingarliðunum þar sem þeir eru fleiri en einn. Það hefur líka komið á daginn að íslenskar skýringar eru víða nauðsynlegar til að þýðandinn hafi betri tilfinningu fyrir því sem uppflettiorðið felur í sér og geta jafnvel komið í veg fyrir misskilning, þar sem ritstjórar íslenska hlutans og markmálanna eru ekki þeir sömu.

grein	no kvk
→ BEYGING	
1	(= trjágrein) langur, trékenndur útvöxtur á tré, trjágrein
2	(= ritsmið) ritsmið, lítil ritgerð, einkum í blaði eða tímariti, sem fjallar um tiltekið efni <i>grein um bókmenntir</i>
3	(= námsgrein) nám, kennsla og fræði um tiltekið efni, námsgrein, fag <i>hún var hæst í þremur greinum á lokaprófinu</i>
4	(= lagagrein o.p.h.) eitt atriði í lista með samfelldum texta, t.d. í reglugerðum og lögum (lagagrein) <i>20. grein fjallar um tollamál</i>
<hr/>	
gera grein fyrir <málinu> útskýra <málið>	
gera sér grein fyrir <ástandinu> skilja og þekkja ástandið	
taka <afsökunina> til greina samþykka <afsökunina>	

Mynd 1. Uppflettiorðið grein. Búið er að vinna orðið á íslensku en engar þýðingar eru komnar inn. Flettan er í fjórum merkingarliðum og eru íslenskar skýringar við þá alla. Notkunardæmi eru í lið 2, 3 og 4. Fyrir neðan strikið eru orðasambönd sem einnig eru skýrð á íslensku.

Ekki er farið eftir stafrófsröð í vinnu við flettugreinarnar heldur er að langmestu leyti unnið eftir merkingarflokkum (því er nánar lýst í kafla 5.6). Stærstu orðflokkarnir þrír, nafnorð, lýsingarorð og sagnir, hafa nær eingöngu verið unnir á þann hátt og auk þess stór hluti atviksorða. Þannig hafa t.d. öll orð innan merkingarflokksins DÝR eða LÖGFRÆÐI verið unnin í einu. Með því að vinna eftir merkingarflokkum fæst mun betra samræmi í framsetningu orðabókarinnar.² Aðrir orðflokkar, fornöfn, forsetningar og samtengingar (þ.e. kerfisorð eða málfræðileg orð), hafa þó að mestu verið unnir sem ein heild, enda eru þeir orðflokkar tiltölulega litlir.

Af öðrum verkþáttum má nefna umræður og samvinnu um flókin og vandmeðfarin orð og gagnrýninn yfirllestur á notkunardæmum sem aðrir ritstjórar hafa búið til. Loks hefur ætíð þurft að sinna samskiptum við samstarfsmennina á hinum Norðurlöndunum, bregðast við spurningum og ábendingum og leysa úr ýmsum málum sem upp hafa komið.

2 Í eldri útgáfum af *Íslenskri orðabók* (t.d. 2. útg. 1993) er víða afar mikið ósamræmi í skýringum á merkingarskyldum flettum þar sem langt er á milli þeirra í stafrófinu, sbr. t.d. skýringarnar á orðunum *harðsjóða* og *linsjóða*.

4.2 Verklag í Skandinavíu

Fjórar ritstjórnir markmálanna starfa við ISLEX-orðabókina í Skandinavíu. Starf þeirra er einkum fólgið í því að þýða íslenska efnid (uppflettiorð, orðastæður, orðasambönd og notkunardæmi) yfir á sitt eigið mál, dönsku, norsku (bókmál og nýnorsku), sænsku og færeysku, og bæta við frekari upplýsingum eða skýringum þar sem þörf krefur. Í Bergen þýðir sami ritstjórinn ætíð bæði á norskt bókmál og nýnorsku. Eins og ritstjórnirnar við íslenska hluta verkefnisins hafa kollegar þeirra í Skandinavíu tekið mið af merkingarflokkunum í þýðingum sínum. Þetta fyrirkomulag tryggir betra samræmi í frágangi greinanna en einnig fær þýðandinn víðari yfirsýn yfir úrval orða innan tiltekins merkingarsviðs sem hann á kost á að nota þegar hann þarf að velja sem nákvæmast jafnheiti við íslenska orðið.

Þótt ekki sé unnið eftir stafrófsröð geta þýðendur hvers tungumáls ávallt séð á hvaða stigi tiltekið uppflettiorð er í vinnslunni þar sem merkt er í sérstakan reit (sem ber vinnuheitið „stig“) þegar búið er að vinna flettugreinina. Ef orð er t.d. fullþýtt á dönsku er sett merki um það í reitinn og þurfa þá ritstjórar danska hlutans ekki að koma aftur að því orði í bráð.

Ritstjórar markmálanna hafa tekið virkan þátt í uppbyggingu ISLEX-orðabókarinnar og veitt ritstjórninni á Íslandi mikilvægt aðhald og er þar því um að ræða ákveðið gæðaeftirlit. Þýðendur hafa m.a. oft komið með tillögur að nýjum uppflettiorðum eða merkingarliðum, þeir hafa óskað eftir að sett yrði inn notkunardæmi við ákveðin orð og lagt til breytingar á dæmum sem þeir hafa talið ófullnægjandi. Einnig hafa þeir komið auga á efnislegar og kerfislegar misfellur, bæði hverjir hjá öðrum og í íslenska hluta verksins.

4.3 Samskipti

Þar sem unnið er við ISLEX-orðabókina í mörgum löndum samtímis er algerlega nauðsynlegt að samskipti milli manna geti farið fram á skjótan og þægilegan hátt. Nú eru til ótal samskiptaleiðir um rafræna miðla svo að það hefur ekki háð verkefninu þótt langt sé á milli þátttakenda landfræðilega séð. Dagleg samskipti hafa að miklu leyti farið fram um tölvupóst en einnig er notast við sérstakt skilaboðakerfi sem er hluti af vinnuumhverfi gagnagrunnsins. Í tengslum við ISLEX-verkefnið var auk þess búið til sérstök lokuð vefsíða (forum-síða) þar sem birtar eru upplýsingar um merkingarflokka sem eru tilbúnir frá

íslensku ritstjórninni og sem þýðendur geta þá byrjað að þýða, og einnig eru þar ýmsar töflur, fundargerðir o.þ.h. og annað sem tengist verkefninu.

grein no kvk	
→	BEYGING
1	(= trjágrein) grein
2	(= ritsmöl) artikkel grein um bókmenntir ein artikkel om litteratur
3	(= námsgrein) (undervisnings)fang, studieretning hún var hæst í þremur greinum á lokaprófinu ho fekk best resultat i tre fag ved avsluttande eksamen
4	(= lagagrein o.þ.h.) paragraf 20. grein fjallar um tollamál paragraf 20 handlar om tollspørsmål
gera grein fyrir <málinu> gjera reie for <saka> gera sér grein fyrir <ástandinu> vera klar over <situasjonen> taka <afsökunina> til greina akseptera <orsakinga> vera á grænni grein sitja på den grøne grein	

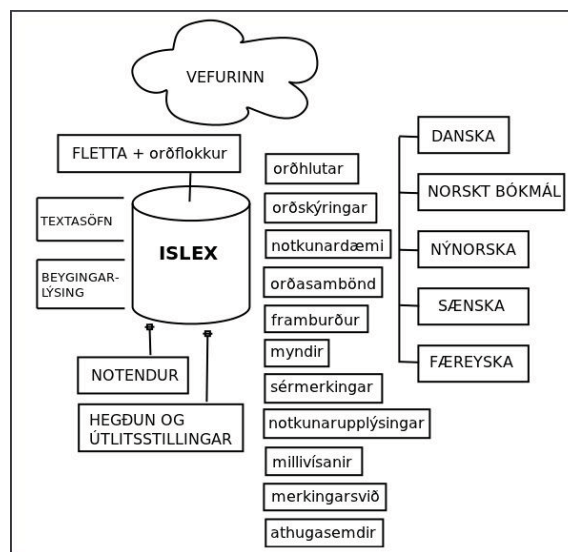
Mynd 2. Uppflettið grein ásamt þýðingum á nýnorsku. Þetta er það sem notandinn sér þegar nýnorsk orðabók er valin.

Ásamt þessum rafrænu samskiptaleiðum hafa að jafnaði verið haldnir tveir þriggja daga vinnufundir ár hvert þar sem allir þátttakendur hafa hist í eigin persónu. Þessir fundir hafa gagnast til upplýsingagjafar, umræðna og áætlana og eru menn sammála um að samskipti um tölvur geta ekki að öllu leyti komið í stað persónulegra samtala. Fundirnir hafa því verið mjög mikilvægur þáttur í samstarfinu. Þýðendur verkefnisins í Skandinavíu hafa einnig hist reglulega til að bera saman bækur sínar.

Eftir að orðabókin var opnuð á vefnum 2011 hafa samskipti við notendur verið hluti af ritstjórnarstarfinu. Á vefsíðu ISLEX er sérstakt kerfi fyrir athugasemdir notenda (undir tenglinum 'Hafa samband') sem hægt er að koma á framfæri við ritstjórnina og hafa margir nýtt sér það. Sagt hefur verið frá villum og öðrum ágöllum, bent hefur verið á orð sem vantar o.s.frv. Allar slíkar ábendingar eru skoðaðar vandlega.

5 Efniviður ISLEX

ISLEX er meðalstór orðabók að því er varðar fjölda uppflettiorða, með tæplega 50.000 orð. Til samanburðar er fjöldi uppflettiorða í *Íslenskri orðabók* (4. útg. 2007) um 94.000, en sú bók telst vera fremur stór. Í gagnagrunni ISLEX eru fólgin margvísleg gögn sem öll eru nákvæmlega skilgreind. Mynd 3 er e.k. flæðirit sem sýnir helstu þætti sem snerta verkefnið.



Mynd 3. Skipan ISLEX-orðabókarinnar. Gagnagrunnurinn er táknaður með stóra sívalningnum. Fyrir ofan hann er reitur með flettu og orðflokki, sem gegna veigamiklu hlutverki í gagnagrunninum. Til vinstri eru mikilvæg gagnasöfn, textasöfnin og beygingar orða. Hægra megin eru reitir með helstu gagnaeyingum í ISLEX, og lengst til hægri eru markmálin, danska, norska, sænska og færeyska. Neðan við ISLEX eru töflur sem geyma upplýsingar um notendur, og stillingar á hegðun og útliti gagnategunda.

Á þeim tíma sem liðinn er frá því að vinna hófst við ISLEX hafa safnast saman miklar upplýsingar í gagnagrunninum, bæði um íslenska orðaforðann og markmálin fimm. Afar mikill efniviður er fólgin inni í sjálfum flettugreinunum, þ.e. því efni sem birt er undir hverri einstakri flettu. Allt efnið byggist á kerfislega flokkuðum upplýsingum ("tegundum"), jafnt íslenski hlutinn og markmálin, og eru t.d. notkunardæmi ein slík tegund og millivísun önnur tegund. Einingarnar hafa sérstakt heiti eftir því hlutverki sem þær gegna í orðabókinni. Alls eru skilgreindar um 80 tegundir gagna, þar af er um helmingurinn

ætlaður fyrir upplýsingar um íslensku auk kerfislegra atriða, en hinn helmingurinn er fyrir ýmis gögn er varða markmálin fimm. Þar eru sjálfar þýðingarnar vitaskuld fyrirferðarmestar en einnig eru þar notkunarupplýsingar af ýmsu tagi og beygingarendingar jafnheita, allt sundurgreint eftir tungumálum.

5.1 Textasöfn og önnur hjálpargögn

Rafræn textasöfn hafa á undanförunum áratugum leyst af hólmi hefðbundna orðtöku sem öll viðamikil orðabókarvinna hefur óhjákvæmilega útheimt í gegnum tíðina, þar sem tilgangurinn hefur verið að leita uppi sem mest af heimildum um tiltekið orð, s.s. merkingu þess, notkun, aldur, útbreiðslu, beygingu o.s.frv. (Um kosti og galla textasafna við orðabókargerð má t.d. benda á Svensén 2004:56-59).

Vitnisburður textasafna hefur fjölþætt gildi í orðabókarvinnu:

- Textarnir gefa vísbendingu um hversu algengt eða sjaldgæft orðið er í málinu.
- Upplýsingar fást um aldursdreifingu orðsins og notkunarvið þess.
- Hægt er að ákvarða merkingu orðs (stundum fleiri en eina) út frá samhengi þess í setningum.
- Ákveðin mynstur orða geta komið í ljós í röðuðum niðurstöðum textasafnsleitar, þ.e. mynstur um fasta orðanotkun sem gera þarf skil í orðabók en auðvelt væri að yfirsjáast með öðrum hætti.
- Fjölbreyttir textar eru mikilvæg hjálp þegar búa þarf til notkunaræmi um orðin.

ISLEX telst til þeirra orðabóka sem eru að miklu leyti grundvallaðar á textasöfnum, og búum við svo vel að hafa aðgang að fjölbreytilegum, rafrænum íslenskum textum. Einkum hafa tvö stór textasöfn verið mikilvæg hjálpartæki í starfinu.

Um svipað leyti og starfið við ISLEX hófst var *Íslensku textasafni* komið fyrir á vef stofnunarinnar. Textasafnið hefur aukist jafnt og þétt með árunum og er stærð þess um mitt ár 2012 um 67 milljónir lesmálsorða. Það hefur að geyma efnisflokkada og að nokkru leyti aldursflokkada texta þar sem leitað er að stökum orðum, í einum efnisflokki eða fleirum í senn. Efnisflokkar textasafnsins eru t.d. SKÁLDRIT, FORNRIT, SAGA OG HEIMSPEKI, LÍFFRÆÐI, RAUNGREINAR OG MATREIÐSLA. Mynd 4 sýnir búið af niðurstöðu leitar í *Íslensku textasafni*.

og breitt bros á tali við tvo því að hún var svo upprifin og yssu í beltisstæð. Norsararnir þir Guðrún Hauksdóttir, fallæg, áttu það sameiginlegt að vera í veitingabúð Gerðubergs. Þaðú Gerðubergs. "Kátir dagar, þara á eftirlaunum. Karlakórinn þangað að kattholti. Karlakórinn þetta: Annar þeirra snaraði sér öll hjónaefni jöfn; faru fyrst þó við verum að spjalla saman í indirfyrirsögn: - UM SJÖHUNDRUÐ einn daginn er hann syngjandi var frændi hans Þórður kakali degi, þegar allir voru nú bara hafði svo sem ástæða til að vera ír og segir að heilsa þeim, allir ír. Þaður verður þaði betri og ír þessi tíðindi. Eg varð feiti í sem ég lék í Stúðnum, feita, egrir en Lüðvík. Lüðvík er ekki svoglumelgi á leiðinni en allir in Óskarsdóttir voru hressar og únna og starfsfólk, "segir Bárá	káta sjólíða. Tvar vörpulegar löggur gengu fyrir gluggann með hendur fyrir aftan kát. Reyndi að lata alls ekki á því bera að mér var verulega illa brugðið. Hann kátu gáfust upp á að biða eftir afgreiðslu og fóru út. Eg horfði á eftir þeim þa kátir og hress Kópavogsmér sem innan tíðar gengur í Kristinna manna tölu. Draumur kátir og yndislegir og greinilega vanir mannanna meðhöndlun. Eg hef einnig heyrt kátir dagar, kátt fólk": Föstudaginn 12. okt. kl. 19 haustfagnaður á Hótel Sögu. kátt fólk": Föstudaginn 12. okt. kl. 19 haustfagnaður á Hótel Sögu, miðar til að kátir Karlár. Hefur áfingar þrjúdaginn 17. sept kl. 13 í Felags- og þjónustumi kátir Karlár hefur áfingar þrjúdaginn 17. sept. kl. 13 í Felags- og þjónustumi kátur inn á sýslumannsskrifstofu og sagði við afgreiðslukonuna: Jája, heldurðu á kát á sýslumannsskrifstofu og svo - vildu þau helga sambandið - í kirkju. Að min kátum kunningsjahópi. Mér dettur í hug saga sem mér finnst passa vel inn í umráðu kátir framhaldsskólameyndur mattu í Fjölbrotaskóla Suðurnesja í gar, en þá var kátur og með mörg járn í eldinum en næsta dag togast ekki upp úr honum orð og ha kátur, en þáir menn verða skemmtir sem hann með vini - lét hann ganga sögur ten kátir, og gleðlegir vindar blésu. Hún sagði við Eyjólf: -Það er vegna þess að ha kátur því að hann er sannanlega gleðiefni giftingardagur dótturinnar, þannig er kátara . Sighvatur það hann þá kveða kveðið það er hann hafði ort um Ástríði. Ött kátari . Sighvatur það hann kveða sér kveði það er hann hafði ort um Ástríði drot kátasti . -Það þýðir lítið fyrir Erlend að þrætta við mig um jarðir, sagði hann. kátur, karlinn, sagði biskupsveinininn. -Hann hefur lítið til að ketast yfir, þa kátir sem ná undir sig jörðum og skreið. Aularnir munu sitja volandi yfir aulask kátur á Ara. -Það er ekki svo lítið sem hún og börnin hafa þjáðst vegna skáldska kátir, brún mín léttist og lundin lifnaði. Margir komu í bílinn á Aflagrandanum, kátari að vera með þessu góða fólk. Þökk sé sóknarnefnd Neskirku fyrir að bjóð káti strákurinn, feiti trúðurinn, og það var einmitt hann sem ég lék í Stúðnum, káta strákurinn, karlvisi þrakkarann, trúðurinn. Þannig sló ég í gegn. Sem ég, einn kátur, hann er ekki skemmtilegur, en hann er ekki villaus. Lüðvík er langfæðas kátir og hressir. Þetta var fjölskylduboð, alveg ágætt. Frændsystkini mín eru al kátur, en Guðrún kennir ballett fyrir skólann og Lovisa jazz og leikfimi. Margir kát sem kennt hefur jazzballett og líkamsrækt í tugi ára. «Við erum búin að stak	(mann) (mann) (mann) (mogg-bladauki) (mogg-brjæftilbl) (mogg-dagbok) (mogg-dagbok) (mogg-dagbok) (mogg-dagbok) (mogg-lesbok4) (mogg-lesbok4) (mogg-lesbok4) (mogg-lesbok4) (mogg-sudurnes) (ofsl) (ofsl) (ofsl) (ofsl) (ottav_fl) (ottav_te) (oxinogjordin) (oxinogjordin) (oxinogjordin) (oxinogjordin) (oxinogjordin) (prestab) (prestab) (riklarhugans) (riklarhugans) (riklarhugans) (riklarhugans) (sedogheyr2006) (sedogheyr2006)
--	---	---

Mynd 4. *Hluti af niðurstöðu leitar að orðinu kátur í Íslensku textasafni. Orðið finnst í öllum beygingarmyndum sínum. Þarna er niðurstöðum raðað eftir textaskrá. Aftast eru tenglar í upplýsingar um viðkomandi ritverk.*

Nokkru áður hafði verið opnaður aðgangur að vefnum *Tímarit.is* sem var samstarfsverkefni Landsbókasafns Íslands - Háskólabókasafns og bókasafna í Færeyjum og á Grænlandi. Tímaritavefurinn er stafrænt safn sem veitir aðgang að milljónum myndaðra eða skannaðra blaðsíðna úr blöðum og tímaritum gefnum út í þessum löndum, en þar er íslenska efnið langfyrirferðarmest. Blöðin hafa að geyma almennt fréttafni, auglýsingar, efni á sviði bókmennta, sagnfræði, ættfræði, þjóðlífs, menningar, atvinnuvega og viðskipta (sbr. <http://timarit.is> – *Um vefinn*). Með árunum hefur leitaraðgangurinn orðið greiðari og blaðsíðum fjölgað, og hefur mikið verið stuðst við þetta einstæða gagnasafn við uppbyggingu ISLEX-orðabókarinnar. Mynd 5 sýnir nút af leitarniðurstöðum á *Tímarit.is*.

Þessi tvö textasöfn eru ómetanlegar ritheimildir um íslensku og hafa þau reynst ómissandi við orðabókarstarfið. Nú er að vísu hægt að leita að orðum á öllum veraldarvefnum og fá fjölmargar niðurstöður úr þeirri leit, en oft er samhengið sem þannig fæst mjög ófullnægjandi heimild um orðið og mikið er um síendurtekin tilvik. Að auki koma mörg orð afar sjaldan fyrir í því efni sem birtist almennt á vefnum og þarf að leita í verk rithöfunda eða fræðimanna til að finna þau. Mörg slík verk er einmitt að finna í *Íslensku textasafni*.

En þrátt fyrir mikla kosti textasafna eins og þeirra sem hér hefur verið lýst hafa þau sín takmörk og ekki er hægt að sækja í þau allan fróðleik. Ef afla þarf ákveðinnar vitneskju um sértækan orðaforða er venjulega betra að leita hennar annars staðar. Með þessu er átt



Mynd 5. Hluti af niðurstöðu leitar að orðinu kátur á vefnum Tímarit.is. Hægt er að smella á greinarnar og sjá stærra samhengi orðanna.

við t.d. latnesk tegundarheiti fugla og skilgreiningar á hugtökum eðlisfræðinnar. Við vinnslu orðabókarinnar hefur því stundum verið stuðst við handbækur og orðabækur um ýmis efni, hinar og þessar vefsíður, og auk þess þekkingu fróðra manna á SÁ og annars staðar. Loks má nefna *Ritmálssafn Orðabókar Háskólans* sem er heljarmikið gagnasafn um íslenskt mál síðustu fimm alda. Það hefur komið að góðum notum í orðabókarstarfi hér á landi síðustu áratugina og hefur líka verið notað við vinnuna við ISLEX.

5.2 Uppflettiordin

Eins og fyrr greinir er ISLEX meðalstór orðabók. Hlutfall orðflokka innan orðaforðans er á þá leið að nafnorð eru í yfirgnæfandi meirihluta, um 68% orðaforðans. Lýsingarorð eru um 17% orðaforðans, sagnir um 9%, atviksorð um 4% og orð af öðrum orðflokkum eru miklu færri. Orðaforðinn skiptist eftir orðflokkum samkvæmt töflu 1:

Orðflokkur	Fjöldi orða	Hlutfall orðaforðans
Nafnorð	33.136	68%
Lýsingarorð	8362	17%
Sagnir	4234	9%
Atviksorð	1815	4%
Töluorð	71	< 1%
Fornöfn	126	< 1%
Forsetningar	235	< 1%
Samtengingar	93	< 1%
Upphrópanir	62	< 1%
Forliðir	136	< 1%
Greinir	1	< 1%
Alls	48.432	100%

Tafla 1. Skipting orða í ISLEX eftir orðflokkum.

5.2.1 Orðaforðinn: val og viðbætur

Eins og greint var frá í kafla 2 var settur saman sérstakur orðabókarstofn hjá Orðabók Háskólans á síðasta áratug 20. aldar (nefndur *norræni orðabókarstofninn*) og varð hann grundvöllur orðaforðans í ISLEX. Í upphaflega orðabókarstofninum er kjarni orðaforðans: algengustu nafnorð, lýsingarorð og sagnir og öll kerfisorð. Þar er áhersla lögð á nútímaíslensku en líka er dálítill orðaforði úr forn máli, eldra máli og skáldamáli. Einnig er nokkuð um fagorð af ýmsum sviðum s.s. líffræði, lögfræði og viðskiptum.

Á þeim árum sem orðabókin hefur verið í smíðum hefur ýmsu verið bætt við upprunalega flettiorðalistann og annað hefur verið fellt brott. Frá því að verkið hófst standa eftir um 41.600 orð úr upphaflega norræna orðabókarstofninum (um 86% orðaforða ISLEX) en um 7.000 orðum hefur verið bætt við (sem svarar til um 14% af orðaforðanum).

Nýrri orðin endurspeglar alloft aðstæður og umræðu líðandi stundar en ekki er það þó algilt. Dæmi um orð sem bæst hafa við eru nafnorðin *auglýsingaherferð*, *áfallahjálp*, *átröskun*, *bílsprengja*, *bökunarplata*, *dagmóðir*, *eftirlitsmyndavél*, *flisteppi* og *hnattvæðing*; lýsingarorðin *alsaklaus*, *áralangur*, *bensínknúinn*, *eldspúandi*, *enskuskotinn*, *gleraugnalaus*, *grænblár*, *heimaprjónaður* og *hjartaveill*; og sagnirnar *aftengja*, *djúpstekja*, *endurlifa*, *fúaverja* og *strádrepa*. Flettulistinn í ISLEX er þannig talsvert sveigjanlegur og opinn fyrir viðbótum.

Um viðbótarorð hafa einkum gilt þessi sjónarmið:

- Orðið er algengt í málinu og hafði ekki ratað inn í upphaflega orðabókarstofninn.
- Orðið er í umræðu líðandi stundar og líklegt virðist að það eigi sér framtíð í tungumálinu (*efnahagskreppa, samkeppnisstaða*). Þetta getur þó stundum verið erfitt að meta.
- Orðið heyrir eða kemur fyrir víða í umhverfinu þótt það sjáist ekki endilega mikið á prenti (*pressukanna, mátunarklefi, slökunartónlist*).
- Ýmis svið orðaforðans hafa verið aukin með orðtöku úr tiltækum heimildum, t.d. tegundir fugla og plantna, landaheiti og fatnaður. Einnig hafa nokkur fagorðasvið verið aukin, t.d. orð sem tilheyra jarðfræði (*kvikuþró* og *megineldstöð* hafa t.d. bæst við þann flokk).

Annars er ekki að öllu leyti auðvelt að skilgreina reglur um val orða í orðabók. Það hefur komið til tals að útbúa tíðniraðaða orðalista úr rafrænum textum og ákveða að “toppurinn” skuli verða uppflettiorð og hefur það raunar verið prófað lauslega í tengslum við ISLEX-verkefnið. Það er þó ekki einfalt verk þegar á hólminn er komið og margar hindranir í veginum. Þættir sem valda örðugleikum eru t.d. röng lemmun (eða vöntun á lemmun), mjög mikið af óviðkomandi orðum s.s. nöfnum, örnefnum, erlendum orðum, tölustöfum og táknum, og eðlislægar takmarkanir textasafna hvað snertir orðaforða málsins. Þegar upp er staðið er það því að miklu leyti máltilfinning ritstjóranna sem ræður vali orðanna, og er stuðnings gjarnan leitað í tiltækum heimildum.

5.2.2 Fleiryrtar flettur

Sum uppflettiorðin í ISLEX-orðabókinni eru fleiryrtar flettur, þ.e. tvö orð eða fleiri. Þar eru atviksliðaflettur áberandi (*að fullu og öllu, alla jafna, af og til, sitt á hvað, smátt og smátt, sums staðar*) en einnig forsetningar (*á milli, fram eftir, í kringum, niður í, upp úr, út undan*), samtengingar (*af því að, né heldur, til þess að*) og fornöfn (*annar hver, sinn hvor*). Að auki er að finna meðal fleiryrtra flettna nokkra nafnliði (*grafísk hönnun, háls-, nef- og eyrnalæknir, heitur pottur*) og lýsingarorðsliði (*annars hugar, frá sér numinn, önnum kafinn*).

Fleiryrtar flettur hafa fengið orðflokksmerkingu í samræmi við setningarstöðu þeirra og hlutverk. Atviksliðafletturarnar eru sambönd tveggja eða fleiri orða sem mynda merkingarlega heild og hafa sömu

stöðu og atviksorð í setningum. Þau fá orðflokksmerkinguna *ao*. Sama gildir um fleiryrtta nafnliði (nefndir *no*) og lýsingarorðsliði (*lo*), en fleiryrtar samtengingar eiga sér lengri hefð í málinu og jafnvel samsettar forsetningar.

Í ISLEX eru alls um 860 uppflettiorð fleiryrtar einingar, og er með þessu farið nokkuð út fyrir hefðina í orðabókum. Hugmyndina að því að taka með svo margar fleiryrtar flettur í ISLEX má að miklu leyti rekja til verkefnisins *Íslenskt orðanet* (www.ordanet.is), sem unnið hefur verið að hjá SÁ frá árinu 2004 undir stjórn Jóns Hilmars Jónssonar. Flest fleiryrttra atviksorða í ISLEX eiga uppruna sinn í því verkefni (sjá Jón Hilmar Jónsson 2009:165–170 og 2012).

Nokkur rök eru með því að hafa fleiryrt uppflettiorð í orðabókinni:

- Fleiryrt sambönd geta oft staðið jafnfætis stökum orðum sem sjálfstæðar, merkingarbærar einingar. Í tvímála samhengi vill svo til að margar fleiryrtar einingar eiga sér einyrt jafnheiti á öðrum málum. Dæmi um það er atviksorðið *allt í einu* sem á sér danska jafnheitið *pludselig* og nýnorska jafnheitið *plutseleg*; að *innanverðu* á sænska jafnheitið *invändigt* og *aftan frá* samsvarar sænsku *bakifrån*. (Þetta á raunar ekki aðeins við milli tungumála heldur á t.d. þríyrta íslenska sambandið *allt í einu* sér einyrta samheitið *skyndilega*.)
- Mörg slík sambönd eru ákaflega algeng í málinu. Aðgangur að þeim í hefðbundnum orðabókum er oft óskýr og ógreiður og þau vilja oft týnast inni í greinunum. Það er þægilegt fyrir notandann að geta gengið beint að þeim.
- Sum orð eru ýmist rituð í einu orði eða tveimur: *ýmiss konar* eða *ýmiskonar*, *sums staðar* eða *sumstaðar*, *báðum megin* eða *báðumegin*. Spyrja má hvort eðlilegt sé að stafsetning ráði því hvernig orði er komið fyrir í orðabók og með hvaða orðflokksmerkimiða. Í ISLEX eru tekin með bæði ritform slíkra uppflettiorða og vísað á milli, en orðflokkurinn er hafður sá sami hvort sem rituð eru eitt orð eða tvö. Þannig eru bæði *sums staðar* og *sumstaðar* greind sem atviksorð.

Við þetta má bæta að orðasambönd í ISLEX-orðabókinni þar sem kjarninn er nafnorð (*fara á fætur*, *hafa hjartað á réttum stað*, *taka til hendinni*) eða lýsingarorð (*fá sér einn gráan*, *vera harður af sér*) eru að jafnaði tilfærð undir kjarnaorðunum en ekki sem sjálfstæðar flettur. Að þessu leyti er farin leið hefðbundinna orðabóka.

5.3 Orðasambönd og orðastæður

Orðasambönd eru áberandi hluti af ISLEX-orðabókinni og er lögð meiri áhersla á að lýsa merkingu þeirra og notkun en hingað til hefur tíðkast í íslensk-erlendum orðabókum. Ástæður þess eru einkum þær að orðasambönd eru mjög stór hluti af tungumálinu og að mörg föst orðasambönd geta verið ógegnisæ og erfið viðureignar fyrir erlenda notendur. Í ISLEX eru öll orðasamböndin þýdd á markmálin og hefur það hagnýtt gildi í báðar áttir í tveggja mála (eða margmála) orðabók þar sem íslenskir notendur geta notfært sér þennan þátt ekki síður en hinir erlendu.

Hugtakið *orðasamband* hefur fremur víða skilgreiningu (sjá t.d. Jón Hilmar Jónsson 2005:xix–xx) en í einfaldasta skilningi er það fast samband tveggja eða fleiri orða. Í ISLEX er gerður greinarmunur á tvenns konar orðasamböndum. *Orðasamband* er notað um orðtak eða myndhverft samband þar sem kjarni þess hefur oft lítið eða ekkert að gera með merkingu sambandsins í heild: <biða> með öndina í hálsinum, <þeir> fara í háar saman, vera á höttunum eftir <upplýsingum>, hafa <lítið> upp úr krafsinu.

Orðastæða samsvarar því sem nefnist á ensku *collocation* og *kollokation* á dönsku og sænsku. Í slíkum samböndum kemur orðið fyrir sem kjarnaliður með föstum fylgdarorðum, einu eða fleirum (*franskar kartöflur, komandi kynslóðir, kaffi og með því, sólin er hátt/lágt á lofti*). Stundum er sagnorð kjarnaliður (*brjóta af sér, ráða <fólk> til starfa, vilja <honum> vel*).

Fjöldi myndhverfra orðasambanda í ISLEX er um 4400 og orðastæðna um 9600. Mikill meirihluti beggja gerða þessara sambanda er fenginn úr gagnagrunni *Stóru orðabókarinnar um íslenska málnotkun* (hér eftir skammstafað SOB) (Jón Hilmar Jónsson 2005) með leyfi höfundar, sem raunar annaðist sjálfur val orðasambanda fyrir nafnorðaflettunar í ISLEX. Sum orðasamböndin hafa þó komið inn í verkið eftir öðrum leiðum.

Framsetning orðasambanda er með svipuðum hætti og í SOB (sjá ítarlega greinargerð um það í inngangi bókarinnar; *ibid.*:xix–xxii). Eitt megineinkenni orðasambanda í SOB er að breytilegir liðir eru afmarkaðir með oddklofum: *krús af <öli>, rissa upp <mynd af fugli>, vera um kyrrt <í sveitinni>*. Þetta fyrirkomulag hefur verið tekið upp í ISLEX-orðabókinni með dálitlum breytingum sem til eru komnar vegna ólíks hlutverks þessara tveggja orðabóka. Þar sem öll orðasambönd í ISLEX eru þýdd á markmálin, líka þau orð sem eru umlukin oddklofum, verður að gera þau þannig úr garði að hægt sé með góðu móti að þýða

þau á önnur mál. Breytilegir liðir hafa því víða verið einfaldaðir og eru mörg orðasambönd í ISLEX talsvert styttri og einfaldari en í SOB.

Í ISLEX hafa eftirfarandi nafnorð flest orðasambönd (fjöldi þeirra er á eftir nafnorðinu og í sviganum er dæmi um orðasamband. Orðastæður eru ekki taldar með):

- (1) hönd 75 (*hafa hendur í hári* <þjófanna>)
- (2) fótur 55 (*setja á fót* <fyrirtæki>)
- (3) mál 53 (*vekja máls á* <þessu>)
- (4) orð 43 (*fá orð í eyra*)
- (5) auga 31 (*koma auga á* <snjalla lausn>)
- (6) maður 30 (*vera maður að meiri*)
- (7) vegur 26 (<ferðast> *á eigin vegum*)
- (8) dagur 25 (<þannig var lífið> *forðum daga*)
- (9) líf 24 (*komast lífs af*)
- (10) ráð 24 (*vera ekki með réttu ráði*)

5.4 Notkunardæmi

Mörg notkunardæmi eru í ISLEX-orðabókinni. Þeim er ætlað að sýna uppflöttiorðið í eðlilegu samhengi og einnig að skerpa greinarmun á mismunandi merkingum orðsins, sé um fleiri en einn merkingarlið að ræða. Nafnorðið *meðferð* skiptist t.d. í fjóra liði. Einn liðurinn merkir 'meðhöndlun' og þar er dæmið *leikföngin hafa fengið slæma meðferð*. Annar liður merkir 'læknismeðferð' og þar er dæmið *hún fékk meðferð við sjúkdómnum*.

Þarna er ekki um að ræða raunveruleg ritdæmi með heimildarvísun eins og tíðkast í sumum orðabókum heldur eru þau samin af ritstjórum verksins. Dæmin eru viðkvæm og vandmeðfarin. Oft er auðveldara að láta lengri dæmi hljóma eðlilegar og betur en mjög stutt dæmi, en í orðabók er samt æskilegt að hafa þau eins stutt og hnitmiðuð og kostur er. Þarna þarf að finna milliveg.

Textasöfn (sbr. kafla 4.1) hafa komið að góðum notum við að búa til dæmin enda eru textar þeirra óþrjótandi uppspretta samhengis orða í tungumálinu. Ekki hefur tíðkast að taka heilar setningar úr textum og gera þær óbreyttar að notkunardæmum heldur eru fengnar þaðan hugmyndir að dæmum þegar innblásturs er þörf. Dæmin eru á eðlilegri nútímaíslensku og eru þau þýdd á öll markmálin eins og orðasamböndin.

Dæmi koma fyrir í öllum orðflokkum. Þótt nafnorð séu langtum fleiri en sagnir eru samanlagt fleiri notkunardæmi í sagnagreinunum heldur en í nafnorðagreinunum (í sögnum eru 10.000 dæmi en 9000 í nafnorðunum). Það er vegna þess að sagnir eru þess eðlis að lýsing þeirra í orðabókinni hvílir mun meira á dæmum en lýsing flestra nafnorðanna. Hér er sýnishorn af dæmum í sagnafluttum þar sem merkingarsviðið er HREYFING/GÖNGULAG. Uppflettiorðið er í sviga fyrir aftan:

áin hlykkjast eftir dalnum (*hlykkjast*)
þeir bröltu yfir grjótið í fjörunni (*brölta*)
blaðamaðurinn skálmaði beint að ráðherranum (*skálma*)
hún ranglaði ein um gamla bæjarhlutann (*rangla*)
ég reikaði niður að höfninni (*reika*)
ég skreiddist fram úr rúminu kl. 12 (*skreiddast*)
gamla konan staulaðist niður stigann (*staulast*)
hann stikaði löngum skrefum út ganginn (*stika*)

Í dæmunum er leitast við að sýna orðið í örstuttu textasamhengi sem er lýsandi fyrir notkun þess. Sum dæmin eru aðeins lýsingarorð með nafnorði: *aðkeypt þjónusta*, *auðfenginn gróði*; undir uppflettiorðunum *aðkeyptur* og *auðfenginn*. Í slíkum tilfellum getur munurinn á dæmi og orðastæðu stundum verið óljós.

5.5 Beygingar

ISLEX-orðabókin gefur ekki upp beygingarendingar eða beygingarflokk við uppflettiorðin eins og gjarnan er gert í prentuðum orðabókum, heldur er við hverja flettu tengill í vef verkefnisins *Beygingarlýsing íslensks nútímsmáls* (á <http://bin.arnastofnun.is>). Þá opnast beygingardæmi orðsins í heild sinni. Í ISLEX eru tenglar í beygingarlýsinguna við nær öll nafnorð, lýsingarorð og sagnir.

Í orðabókinni er einnig að finna upplýsingar um beygingarendingar jafnheita markmálanna, og er því lýst í kafla 5.8.3.

5.6 Merkingarflokkar

Öllum orðaforðanum í ISLEX er skipt upp í merkingarflokka. Í köflum 3.1 og 3.2 kom fram að bæði íslensku flettugreinarnar og þýðingarnar hafa verið unnar eftir þessum merkingarflokkum. Þeir eru tæki fyrir

ritstjórnarvinnuna og eru ekki sýnilegir notendum. Flokkarnir bjóða þó upp á ýmsa möguleika til frekari birtingar á orðabókarefninu, t.d. væri með áframhaldandi þróun hægt að nota þá til að útbúa samheita- og skyldheitabálka innan orðaforðans.

Merkingarflokkarnir eru hópar merkingarskyldra orða sem eru allbreytilegir að stærð. Þeir geta verið fáein orð upp í að vera mörg hundruð, en dæmigerður merkingarflokkur er nokkrir tugir orða. Dæmi um merkingarflokka eru DÝR, FATNAÐUR, HÚSBÚNAÐUR, BÍLAR (mest hlutbundin nafnorð); HLJÓÐ, ÓVEÐUR, SAMKOMA, ÞJÓÐARHEITI, STAÐARHEITI, STARFSHEITI, ÆTTARTENGL (óhlutbundnari orð); EFNAFRÆÐI, JARÐFRÆÐI, LÍFFRÆÐI, LÆKNISFRÆÐI, MÁLFRÆÐI, STÆRÐFRÆÐI (sérsvið innan orðaforðans). Áberandi merkingarflokkar meðal lýsingarorða eru ALDUR, ÁNÆGJA, ÁSTAND, DUGLEYSI, FJÁRHAGUR, HEILSA, HOLDAFAR, LITUR, LÖGUN og ÚTLIT. Merkingarflokkar sagna eru t.d. BREYTING, DAUÐI, DVÖL, HREYFING, SNERTING, SVEFN, ÚRKOMA og VINDUR. Merkingarsviðin skarast mjög víða, t.d. tilheyrir orðið *gulrót* bæði flokknum GRAS og MATUR, og *kennari* er bæði í flokkunum STARFSHEITI og SKÓLI.

Til að gefa skýrari mynd af þessu má taka tvö dæmi um orðahópa sem tilheyra sama merkingarflokknum. Í flokknum STJÓRNARFAR eru einkum nafnorð, m.a.

alræði, bananalýðveldi, einræði, einræðisríki, einveldi, flokksveldi, flokksræði, fulltrúalýðræði, harðstjórn, hægristjórn, konungdæmi, lénskipulag, lögregluríki, ógnarstjórn, skrifræði, vinstristjórn, þingkjörinn, þjóðveldi

Í flokknum GANGA/GÖNGULAG eru sagnir mest áberandi, m.a.

arka, feta, ganga, haltra, hökta, kjaða, keifa, labb, labba, léttstígur, læðast, lötra, marsera, paufast, plampa, rangla, ráf, ráfa, ráp, rápa, reika, röltá, silalegur, skálma, skjögra, slaga, spígspora, staulast, stjákl, stjákla, stórstígur, storma, strunsa, tipla, trítla, tölta, þungstígur

Með því að einbeita sér að einum merkingarflokki í einu fær ritstjórinn yfirsýn yfir heilt svið orðaforðans, og hefur það reynst heppilegri aðferð en að vinna fletturnar eftir stafrófsröð.

5.7 Myndir og hljóð

Fjöldinn allur af myndum eru í ISLEX. Þeim er ætlað að vera til prýði

og gleði en þær hafa einnig víðtækara hlutverk. Mynd getur skýrt merkingu sumra orða betur en hægt er að gera með orðum (*blævængur, gleraugu, íkorni, skæri, tannhjól, tvinnakefli, þvottaklemma*), og stundum er hentugt að hafa myndir til að greina í sundur mismunandi merkingar orða, séu þær fleiri en ein. Myndirnar gera auk þess orðabókina meira aðlaðandi í augum notenda. Í ISLEX eru um 3200 ljósmyndir og teikningar úr ýmsum áttum. Þar af eru um 700 myndir eftir Jón Baldur Hlíðberg af fyrirbærum úr ríki náttúrunnar: plöntum, fuglum, fiskum og fleiri dýrum.



Mynd 6. Myndskreytingar með dýrum og plöntum eftir Jón Baldur Hlíðberg.

Framburður uppflöttiorða í ISLEX er gefinn með hljóddæmum en ekki með hljóðritunartáknum. Hægt er að smella á hljóðtengil til þess að hlusta á orðið með eðlilegum íslenskum framburði.³ Sum orð eru þannig að máhljóð þeirra eru breytileg eftir beygingarmynd (t.d. *dagur – degi, saga – sagna, trefill – treflar*), og í slíkum tilvikum er ætlunin að hafa ekki aðeins hljóddæmi með grunnformi orðsins heldur valdar orðmyndir þar sem tilbrigði í framburði koma fram.

5.8 Markmálin

Markmálin í ISLEX-orðabókinni eru danska, norskt bókmál, nýnorska, sænska og færeyska eins og áður segir. Hér verða markmálunum ekki gerð tæmandi skil en aðeins nefnd nokkur atriði (sjá nánar Aldísi Sigurðardóttur *et al.* 2008:784–786 og Margunn Rauset *et al.* 2012).

Í hverju markmálslandi hafa að jafnaði starfað 1–4 ritstjórar, þ.e. þýðendur. ISLEX-orðabókin gerir miklar kröfur til þeirra. Þýðendurnir þurfa að vera mjög vel að sér bæði í íslensku og markmálinu, þeir þurfa að vera næmir á báðar tungurnar og fundvísir á réttu jafnheitin við íslensku flettunar. Bakgrunnur þýðenda ISLEX er mjög mismunandi. Sumir þeirra eru Íslendingar, aðrir Skandínavar, nokkrir eru af blönduðu þjóðerni. Sumir Skandínavanna hafa starfað sem sendikennarar um tíma á Íslandi og kennt móðurmál sitt við Háskóla

³ Þegar þetta er ritað er þessi verkþáttur enn á vinnslustigi.

Íslands og lært íslensku í leiðinni. Aðrir eru Íslendingar sem hafa búið áratugum saman í einhverju öðru Norðurlandanna.

5.8.1 Þýðingarnar: helstu áskoranir

Starf þýðandans er að finna og skrá á sínu tungumáli sem nákvæmust jafnheiti við íslensku uppflettiorðin, í öllum merkingarliðum þegar þeir eru fleiri en einn. Einnig þarf að þýða orðastæður, orðasambönd og notkunardæmi. Víða þarf að bæta frekari upplýsingum við jafnheitin sem valin eru til nánari skýringar.

Engin sérstök regla gildir um það í hvaða röð markmálin koma inn í gagnagrunninn. Stundum geta sænskar þýðingar orðið fyrstar, stundum þær dönsku o.s.frv. Markmálin styðja hvert annað í mörgum tilfellum en þau geta líka haft truflandi áhrif hvert á annað við þýðingarnar. Í vinnuumhverfi gagnagrunnsins blasa öll málin samtímis við ritstjórunum. Þá getur komið sér vel fyrir þýðandann að horfa á notendaviðmót orðabókarinnar, en þar er hægt að sjá eitt tungumál í einu án áreitis frá hinum málunum. Oft finnst þýðendunum gott að koma að hreinni og ósnertri flettu en í öðrum tilvikum geta þeir nýtt sér rannsóknarvinnu annars þýðanda ef hann hefur orðið á undan til þess að þýða t.d. viss fræðiorð á sitt mál. Mörg fræðiorð og alþjóðleg orð eru að stofni til eins á norðurlandamálunum öðrum en íslensku (og e.t.v. færeysku), t.d. er orðið *veðurfræði* nefnt *meteorologi* á dönsku, norsku og sænsku, og *hnattvæðing* er *globalisering* á hinum málunum.

Norrænu tungumálin eru oft sviplík en stundum standast þau ekki á. Það kemur fyrir að sum málin eiga jafnheiti við íslenska orðið á meðan önnur þurfa að grípa til skýringa:

afréttur no kk

Norska: *beite, fjellbeite*

Danska: *græsgang* (i *vildmarken*; *fælles græsning for et områdes fårehold*)

Færeyska: *fjallhagi*

Oft koma upp slík tilfelli þar sem er aðeins í sumum málunum hægt að þýða beint með orði á móti orði. *Afréttur* er fyrirbæri sem þarfnast lengri útskýringar á dönsku, væntanlega vegna þess að landslag og landshættir í Danmörku eru þess eðlis að nákvæmlega sambærilegt hugtak fyrirfinnst ekki í dönsku. Danska þýðingin verður því útskýringarkennd. Aftur á móti er *afréttur* til í Noregi og Færeyjum líkt og á Íslandi og þar getur þýðandinn fært inn jafnheiti á sínu máli.

Sveitungi er orð þar sem jafnheiti er til á norsku en ekki á dönsku og sænsku:

sveitungi no kk

Norska: *sambygding*

Danska: *person som er fra samme område som nogen*

Sænska: *person som är från samma ort eller trakt som angiven person*

Sumt er vandasamara en annað í þýðingum á efninu í ISLEX. Fag-orðaforði íslenskunnar, þ.e. orð sem heyra til sérsviðum, getur verið mjög erfiður viðureignar fyrir þýðendurna og í sumum löndunum hafa þeir fengið aðstoð fræðimanna við þann orðaforða. Sum orð innan sérsviða eins og bragfræði, jarðfræði og lögfræði geta verið illskiljanleg mörgum íslenskum málnotendum og er ekki erfitt að gera sér í hugarlund hve tímafrekt getur verið að þýða þau á önnur mál, sérstaklega þar sem handbækur um þessi fög með erlendum atriðisorðum eru ekki ýkja margar.

Enda þótt leitast hafi verið við að hafa málsnið notkunardæmanna í ISLEX sem hlutlausast hefur einn þýðendanna getið um gildru sem auðvelt er að falla í við þýðingarnar og sem þurfi að varast, en það er að íslenskan virkar oft fremur hátíðleg á Skandínava. Það getur bæði átt við um einstök orð og byggingu setninga, og það þarf að gæta þess að hafa þýðingarnar ekki í of hátíðlegum stíl þar sem það á ekki við.

5.8.2 Upplýsingar um jafnheitin

Þegar tvö eða fleiri erlend jafnheiti eru valin sem þýðing við tiltekið íslenskt orð er oft bætt við fróðleik um rétta notkun hvers jafnheitis. Þetta kemur sér vel fyrir þá málnotendur sem hafa það markmál ekki að móðurmáli. Sænsku þýðingarnar á so. *rymja* og no. *fimmkall* lýsa þessu:

rymja, v: *bröla (om tjur), råma (om ko), skria (om åsna)*

fimmkall, n m: *femma (om myntet); fem spänn (om beloppet)*

Orðið *amma* er ekki til í sænsku í sama skilningi og í íslensku, dönsku og norsku. Í sænsku þarf að taka fram hvort um er að ræða föðurömmu eða móðurömmu:

amma no kvk

Danska: *bedstemor*

Norska: *bestemor*

Sænska: *farmor (móðir föður); mormor (móðir móður)*

Sums staðar er að finna upplýsingar um íslenska uppflettiorðið sem skýra nánar merkingu þess eða notkun fyrir erlendum málhöfum. Eftirfarandi danskar þýðingar sýna þetta:

kvæntur, adj: *gift (om mand)*

skólasystir, n f: *skolekammerat (om pige)*

Íslensku orðin að ofan eru takmörkuð við ákveðið kyn og hafa þannig þrengri merkingu en dönsku jafnheitin *gift* og *skolekammerat*. Í slíkum tilvikum er nauðsynlegt að bæta við frekari skýringu sem er látin birtast innan sviga.

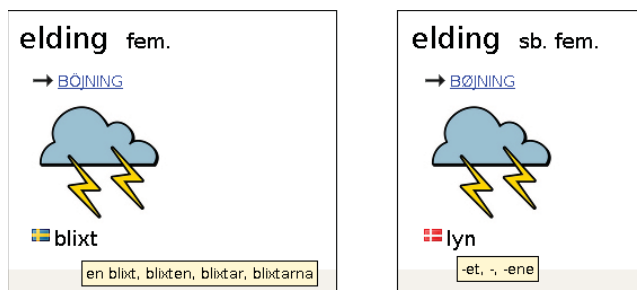
5.8.3 Málfræðiupplýsingar um markmálin

Þegar Íslendingur notar ISLEX-orðabókina getur verið gagnlegt fyrir hann að vita hvort danska jafnheitið er í samkyni eða hvorugkyni eða hvernig þátíð er af sænskri sögn. Þessar upplýsingar er að finna í gagnagrunni orðabókarinnar. Þær eru ekki sýnilegar á yfirborðinu en birtast þegar mús er rennt yfir erlenda jafnheitið, þá sprettur fram rammi með beygingarendingum viðkomandi orðs. Markmálin fara dálítið mismunandi leiðir til að birta slíkar málfræðiupplýsingar eins og sést á næstu fjórum myndum. Myndir 7 og 8 sýna beygingarendingar sænska orðsins *fjäril* og danska orðsins *sommerfugl* 'fiðrildi'.



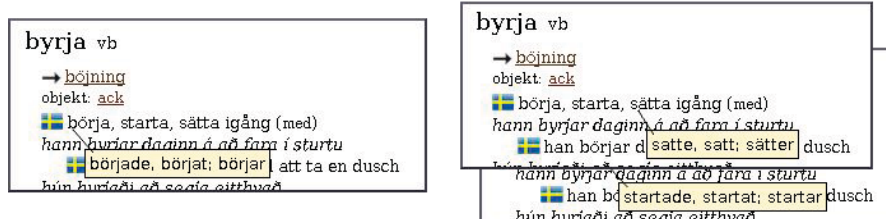
Mynd 7 og 8. Uppflettiorðið fiðrildi með sænskri og danskri þýðingu. Í sænsku gerðinni birtist rammi með beygingarmyndum jafnheitisins þegar mús er rennt yfir það: orðið með óákveðnum greini, með viðskeyttum ákveðnum greini, fleirtala orðsins og fleirtalan með greini. Í dönsku gerðinni birtast beygingarendingar jafnheitisins á sama hátt: viðskeyttur ákveðinn greinir, fleirtöluending og ákveðinn greinir í fleirtölu.

Myndir 9 og 10 sýna beygingarendingar sænska orðsins *blíxt* og danska orðsins *lyn* 'elding'.



Mynd 9 og 10. Uppflettiorðið elding með sænskri og danskri þýðingu. Músinni hefur verið rennt yfir jafnheitin svo að beygingarendingar þeirra koma í ljós.

Málfræðiupplýsingar af þessu tagi eru hafðar við nafnorð, lýsingarorð og sagnir. Mynd 11 sýnir beygingarendingar þriggja sagna í sænsku. Í sérhverju af jafnheitunum þremur eru fólgnar upplýsingar um beygingu þess.



Mynd 11. So. byrja með þýðingum á sænsku. Beygingarendingar sænsku jafnheitanna börja, starta og sätta igång sjást þegar mús er rennt yfir viðkomandi sænskt orð.

Þótt sænskir og danskir notendur þurfi varla á slíkum fróðleik að halda um sín eigin mál geta þær komið sér vel fyrir Íslendinga (og aðra) sem ætla sér að tala eða skrifa sænsku eða dönsku.

6 Notendaaðgangur að ISLEX

Aðgangur að ISLEX-orðabókinni er á vefnum. Það form býður vitanlega upp á aðrar og á margan hátt líflegri aðferðir við framsetningu efnisins en kostur er á í hefðbundinni, prentaðri bók, t.d. hljóddæmi, myndir og hreyfimyndir, svo og tengla til að ferðast á milli flettna. Gefinn er kostur á nokkrum mismunandi leitarmöguleikum. Mynd 12 sýnir orðið *eldgos* eins og það birtist í notendaviðmóti orðabókarinnar.



Mynd 12. Íslenskt notendaviðmót ISLEX. Flettan er eldgos og hér má sjá tengil í beygingu orðsins, orðhlutaskil (eld-gos), mynd af eldgosi og sænska þýðingu orðsins neðst. Dálkurinn til hægri sýnir orðið í stafrófsröð með nærliggjandi orðum fyrir ofan og neðan. Fleiri leitaradgerðir koma í ljós þegar smellt er á Flóknari leit efst t.h..

Þegar notandi flettir upp í ISLEX velur hann sjálfur hvaða orðabók hann vill fá, danska, norska (bókmál), nýnorska eða sænska orðabók. Einnig er hægt að sjá öll málin samtímis. Mynd 13 sýnir orðið *ljósapera* þar sem öll markmálin birtast í einu.



Mynd 13. Orðið *ljósapera* með þýðingum á dönsku, norsku (bókmáli og nýnorsku), sænsku og færeysku. Færeyskan sést bara þegar smellt er á hnappinn allar, en ekki er búið að opna færeysku orðabókina að öðru leyti.

ISLEX-orðabókin hefur fimm mismunandi notendahami, en þá eru umgjörð og textar vefsíðunnar á mismunandi málum: íslensku, dönsku, norsku bókmáli, nýnorsku og sænsku. Verkefnið á sér veffang í hverju þátttökulandi: www.islex.is, www.islex.dk, www.islex.no og www.islex.se. Ef notandi fer t.d. inn á vefsíðuna www.islex.se kemur sjálfkrafa upp sænsk vefsíða orðabókarinnar. Notandinn getur þó ætíð sjálfur stillt tungumál vefsíðunnar.

Þegar leitað er í orðabókinni er slegið inn orð í þar til gerðan reit. Niðurstöður úr leitinni fara að nokkru eftir tungumáli notendahamsins en einföld leit skilar þrenns konar niðurstöðum:

- íslensku uppflettiorði
- orðmynd í íslensku (þ.e. hvaða beygingarmynd sem er af uppflettiorðunum)
- jafnheiti á völdu markmáli

Á mynd 14 hefur verið slegið inn orðið *skál* og vinstra megin á myndinni sjást niðurstöður úr þeirri leit. Í íslenskum notendaham er ætíð leitað að jafnheitum í öllum markmálunum og því sjást margir tungumálafánar við orðin í listanum. Fyrsta orðið í listanum er *skál*, en hún er uppflettiorð og fær því mest gildi í leitinni. Næsta orð er *skal* sem er beygingarmynd af so. *skulu*. Í þriðja sæti listans og niður úr eru jafnheitin *skal*, *skål* og *skäl* á fjórum málum. “Loðin leit” (e. *fuzzy search*) er sjálfgild leitarhegðun. Með því er átt við að leyfð er ákveðin ónákvæmni í innslætti til hagræðis fyrir notandann, og er þannig ekki gerður greinarmunur á bókstöfunun *a*, *á*, *å* og *ä*; *d* og *ð* eru jafngildir stafir í leit o.s.frv.



Mynd 14. Íslenskt notendaviðmót ISLEX. Leit að *skál* skilar mörgum ólíkum niðurstöðum, fyrst uppflettiorðinu *skál*, svo orðmyndinni *skal* (af *skulu*) og því næst ýmsum jafnheitum markmálanna. Hægra megin við tungumálafánana eru íslensku uppflettiorðin (feitletruð) sem jafnheitin koma fyrir í.

Hægt er að leita að skandínávískum orðum sem ekki eru íslensk uppflettiorð en óvíst er hvaða árangur slík leit ber. Mynd 15 sýnir niðurstöðu leitar að orðinu *monster* í sænsku notendaviðmóti. *Monster* er hvorki til sem uppflettiorð né beygingarmynd, heldur eru í niðurstöðulistanum flettur þar sem orðið kemur fyrir sem sænskt jafnheiti. Neðsta orðið *mönster* finnst líka vegna þess að leitin er loðin.



Mynd 15. Sænskt notendaviðmót ISLEX. Leitað hefur verið að monster. Niðurstöður leitarinnar birtast í tveimur dálkum: vinstra megin eru sænsku jafnheitin og hægra megin við þau eru viðkomandi íslensk uppflettiorð. Næsta skref er að velja eitthvert af orðunum úr listanum og fara inn í viðkomandi flettugrein.

Þegar flóknari leit er valin eru fleiri möguleikar í boði. Hægt er að:

- Slökkva á leit að *uppflettiorði* (þá er bara leitað að erlendu jafnheitin).
- Slökkva á leit að *jafnheitin* (þá er aðeins leitað að íslensku uppflettiorði og orðmynd þess).
- Kveikja á leit í *texta*, en þá finnast orð eða stafastrengir sem koma fyrir í notkunardæmum og orðasamböndum, jafnt á íslensku sem markmálunum.
- Slökkva á loðinni leit.
- Tilgreina nánar þau markmál sem leitin nær til með því að haka við eða slökkva á reitunum *danska*, *norska*, *nýnorska* og *sænska*.

ISLEX hefur einnig verið aðlagð að helstu stýrikerfum snjallsíma. Þá er farið á vefsíðuna á venjulegan hátt og sérstakur hugbúnaður skynjar hvaða tæki verið er að nota og sér um að birta vefinn í samræmi við það.



Mynd 16. Flettan eldgos í snjallsíma.

7 Frekari þróun og notkunarmöguleikar

Verkefni eins og ISLEX felur í sér marga spennandi möguleika til frekari þróunar. Hægt er til dæmis að bæta við fleiri tungumálum í gagnagrunninn, og í undirbúningi er að finnska verði sjötta markmál orðabókarinnar. Einnig hafa enska, franska og þýska verið nefndar til sögunnar sem vænlegir kostir.

ISLEX er nú gefin út sem veforðabók en útgáfuformið getur breyst með tímanum. Þegar upphaflega var gengið til samstarfs við sænska og norska aðila stóð til að ISLEX yrði prentaðar orðabækur milli íslensku og sænsku annarsvegur og íslensku og norsku hinsvegar. Það fyrirkomulag getur vel orðið að veruleika en hafa þarf í huga að talsverður kostnaður fylgir umbroti og prentun orðabóka. Ennfremur er þá lokað fyrir möguleikann á viðbótum og leiðréttingum sem svo fyrirhafnarlítið er að bæta inn í rafrænt verk. Á móti kemur að prentuð orðabók hefur ætíð nokkurn híbýlapokka í bókahillu eða á borði og sumir vilja frekar nota orðabækur á pappír en í tölvu.

Með tilkomu rafbóka og lestölva (lesbretta) er orðið til þriðja sniðið sem hvorki er á vef né pappír. Vel má vera að slík útgáfa á ISLEX líti dagsins ljós innan fárra ára. Orðabækur á geisladiskum hafa áður gegnt svipuðu hlutverki en útgáfa efnis á geisladiskum hefur nú að miklu leyti vikið fyrir öðrum miðlum.

ISLEX er líka kjörinn efniviður í nýja íslensk-íslenska orðabók, en slíka bók af hentugri stærð hefur sárlega vantað í nokkra áratugi. Samhliða vinnunni við ISLEX hefur orðið til stofn í nýja íslenska orðabók yfir nútímamálið með vandlega völdum uppflettiorðum, orðskýringum og notkunardæmum. Orðskýringar á íslensku í ISLEX eru upphaflega tilkomnar til þess að auðvelda þýðendunum verk sitt en þær hafa gert það að verkum að vísir að íslensk-íslenskri orðabók hefur myndast sem hliðarafurð. Skýringarnar eru þó ekki alltaf nægilega vandaðar til að vera birtingarhæfar eins og sakir standa. Ekki þyrfti þó mjög mikla viðbótarvinnu til að hægt væri að gefa út handhæga, meðalstóra íslenska orðabók.

Eins og hún er hönnuð hentar ISLEX-orðabókin vel til að nota við tungumálakennslu í skólum þar sem hún er mjög þægilegt verkfæri. Einnig er hún kjörin til þess að búa til kennsluefni fyrir grunnskóla og menntaskóla, og hefur nú þegar verið unnið slíkt verkefni fyrir menntaskóla (sjá www.flensborg.is/islex).

ISLEX er ekki aðeins orðabók heldur hefur sprottið af verkefninu stórt gagnasafn um norrænu tungumálin. Sérhvert íslenskt uppfletti-

orð er tengt jafnheitum sínum á dönsku, norsku, sænsku og færeysku í gegnum færslur í gagnagrunninum. Allur þessi efniviður með vandlega skilgreindum upplýsingum gerir það að verkum að ISLEX er gott hráefni í ýmiss konar máltæknileg verkefni þar sem fengist er við norðurlandamálin. Geta má þess að gert hefur verið samkomulag við forráðamenn META-NORD-verkefnisins um að hagnýta megi ákveðin gögn úr ISLEX. META-NORD er verkefni sem Norðurlöndin og Eystrasaltslöndin standa að, og er markmið þess að efla gagnsöfn um viðkomandi tungumál sem nýst geti í margvísleg máltækniverkefni (sjá nánar á www.meta-nord.eu).

Í vinnunni við verkefnið hefur safnast upp mikill fjöldi jafnheita á markmálunum. Svo dæmi sé tekið eru um 68.000 sænsk jafnheiti í ISLEX (41.000 jafnheiti ef hvert er talið einu sinni). Svipað má segja um hin markmálin. Vegna þessa auðuga orðaforða gæti efnið í ISLEX hentað sem grunnur í orðabækur sem snúa öfugt, þ.e. í sænsk-íslenska orðabók o.s.frv.

Af ofangreindu er ljóst að ISLEX felur í sér margvíslega möguleika til stækkunar, þróunar, útgáfu og rannsóknarverkefna. Tíminn leiðir svo í ljós hvað á eftir að spretta af þessu orðabókarverki.

8 Lokaorð

Hér hefur verið fjallað vítt og breitt um veforðabókina ISLEX en einkum þá þætti sem snúa að íslenska viðfangsmálinu. Fyrst var sagt frá bakgrunni verkefnisins og upphafi þess, og því næst var farið almennum orðum um hina rafrænu orðabók. 4. kafli fjallaði um verklagið við ritstjórn orðabókarinnar, fyrst um íslenska hlutann og síðan um starfið við markmálin í nágrannalöndunum. Í 5. kafla sagði frá efni og uppbyggingu orðabókarinnar og var þar m.a. fjallað um notkun textasafna í starfinu, orðaforðann í ISLEX, uppbyggingu flettugreina, orðasambönd og notkunardæmi. Greint var frá merkingarflokkum sem hafa verið nýttir við að gera flestum flettugreinunum skil, bæði við vinnslu íslenska hlutans og markmálanna. Rætt var um markmálin og helstu áskoranir þar, en hafa ber í huga eins og fyrr er getið að þessi grein er rituð út frá sjónarhóli ritstjóra íslenska hluta ISLEX. Í 6. kafla var lýst notendaviðmóti orðabókarinnar á vefnum og í 7. kafla var fjallað um ýmsa möguleika sem búið geta í svona verki og hugað að framhaldi þess.

Fram kom að orðabókin hefur verið unnin af mörgum ritstjórnnum samtímis í mörgum löndum og hefur íslenska ritstjórnin leitt verkið. Þetta verklag hefur óhjákvæmilega verið nokkuð þungt í vöfum og gert samhæfingu orðabókarstarfsins allkrefjandi. Þrátt fyrir það hefur starfið við ISLEX almennt gengið greiðlega en svo margslungið verkefnið útheimtir talsverða lipurð í samskiptum, bæði á milli landa og innan ritstjórnarhópanna innbyrðis. Einnig hefur þurft að halda mjög vel á spöðunum varðandi nýtingu á sameiginlegum styrkjum sem verkefnið hefur fengið, skipulag starfsins, skipulag funda o.s.frv.

ISLEX-orðabókin var opnuð almenningi þann 16. nóvember 2011, á degi íslenskrar tungu. Viðtökur hennar hafa verið afar jákvæðar og orðabókin fékk fljótt fjölmarga notendur. Verkinu var þó ekki fulllokið þegar það var opnað, t.d. vantaði talsvert upp á að norsku þýðingarnar væru tilbúnar, margar myndir vantaði og upptökur af framburði orðanna voru ekki tilbúnar. Notendur hafa verið þolinmóðir gagnvart þessum annmörkum enda kemur fram fyrirvari um slíkt á vefsíðu orðabókarinnar þar sem segir frá því að efnið sem vantar sé á leiðinni. Margir notendur hafa haft samband og lýst yfir ánægju sinni með verkefnið, sem vonandi auðnast að vaxa áfram og dafna.

Heimildir

Aldís Sigurðardóttir, Anna Hannesdóttir, Halldóra Jónsdóttir, Håkan Jansson, Lars Trap-Jensen, Þórdís Úlfarsdóttir. 2008. ISLEX – An Icelandic-Scandinavian Multilingual Online Dictionary. *Proceedings of the XIII Euralex International Congress*, bls. 779–789. (Á geisladiski).

Beygingarlýsing íslensks nútímsmáls. Stofnun Árna Magnússonar í íslenskum fræðum: <http://bin.arnastofnun.is>.

Halldóra Jónsdóttir og Þórdís Úlfarsdóttir. 2010. ISLEX – Íslandsk-Skandinavisk webordbog. *Málfríður* 2. tbl. 2010: 11–15.

ISLEX-orðabókin. Stofnun Árna Magnússonar í íslenskum fræðum: www.islex.is, www.islex.dk, www.islex.no og www.islex.se.

Íslensk orðabók. 1963. Ritstjóri Árni Böðvarsson. Reykjavík: Bókaútgáfa Menningarssjóðs.

Íslensk orðabók. 1993. 2. útg. aukin og bætt, 7. prentun. Ritstjóri Árni Böðvarsson. Reykjavík: Mál og menning.

Íslensk orðabók. 2007. 4. útg. Ritstjóri Mörður Árnason. Reykjavík: Edda.

Íslenskt orðanet. Stofnun Árna Magnússonar í íslenskum fræðum: www.ordanet.is.

Íslenskt textasafn. Stofnun Árna Magnússonar í íslenskum fræðum: www.lexis.hi.is/corpus/leit.pl.

- Jón Hilmar Jónsson. 2005. *Stóra orðabókin um íslenska málnotkun*. Reykjavík: JPV.
- Jón Hilmar Jónsson. 2009. Lemmatisation of Multi-word Lexical Units: Motivation and benefits. Í: H. Bergenholtz, S. Nielsen og S. Tarp (ritstj). *Lexicography at Crossroads*, bls. 165–194. Bern, Berlin o.v.: Peter Lang.
- Jón Hilmar Jónsson. 2012. Að fanga orðaforðann: orðanet í þágu orðabókar. *Orð og tunga* 14:39–65.
- Kristín Bjarnadóttir. 1998. *Norræna verkefnið*: www.lexis.hi.is/kristinb/norr.pdf.
- Ljunggren, Anders. 2011. *Några skärvor som kan fogas till en bild av vägen fram till ISLEX tillkomst*. <http://islex.is/islex/greinar/Ljunggren-nov2011.pdf>.
- Margunn Rauset, Anna Helga Hannesdóttir og Aldís Sigurðardóttir. 2012. Ein-, to- eller fleirspråklig ordbok? Í: Eaker, B., L. Larsson og A. Mattisson (ritstj.). *Nordiska studier i lexikografi* 11, bls. 512–523. Lundur. *META-NORD-verkefnið*. www.meta-nord.eu.
- Ritmálssafn Orðabókar Háskólans*. Stofnun Árna Magnússonar í íslenskum fræðum: www.arnastofnun.is/page/arnastofnun_gagnasafn_ritmal.
- Svavar Gestsson. 2011. *Ævintýri handa öllum heiminum*. http://islex.is/islex/greinar/avarp_SG-nov11.pdf.
- Svensén, Bo. 2004. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedts Akademiska Förlag.
- Tímarit.is*. Landsbókasafn Íslands – Háskólabókasafn: <http://tímarit.is>.
- Þórdís Úlfarsdóttir. 2011. *Upphaf ISLEX-verkefnisins*. http://islex.is/islex/greinar/Islex_upphaf.pdf.

Lykilorð

tívmála orðabók, margmála orðabók, veforðabók

Keywords

bilingual dictionary, multi-lingual dictionary, online dictionary

Abstract

The Icelandic-Scandinavian dictionary ISLEX was opened on the web in November 2011. The outcome of the project is twofold. Firstly, a long awaited Icelandic-Scandinavian dictionary has been produced using the web-medium and its many possibilities. Secondly, the project has resulted in the compilation of a large amount of data on the Nordic languages, and provided important experience in jointly building such work in five countries simultaneously, as the Icelandic source language has been processed in Reykjavík and the target languages in Copenhagen, Bergen, Göteborg and Tórshavn, respectively. ISLEX has thus contributed to Nordic language collaboration.

The ISLEX-dictionary describes modern Icelandic, including a range of phrases, collocations and examples of use, all of which are translated into the five TLs, Danish, Norwegian (Bokmål og Nynorsk), Swedish and Faroese. ISLEX is thus a multi-lingual dictionary within which are contained five bilingual dictionaries. The electronic form offers direct search both in Icelandic and the TLs and is therefore ideal for comparative studies between the languages. The project also offers other possibilities for research. The material is a good base for various internordic projects within education and language technology. Moreover, the whole Icelandic vocabulary is semantically classified, a classification which extends to the TLs too up to a point.

*Þórdís Úlfarsdóttir
Stofnun Árna Magnússonar í íslenskum fræðum, orðfræðisvið
Neshaga 16, 107 Reykjavík
dlsa@hi.is*

Rósa Elín Davíðsdóttir

Hlutverk tvímála orðabóka

Ólíkar notendaparfir

í íslensk-frönsku ljósi

1 Inngangur

Frönskukennsla á sér langa sögu á Íslandi en franska hefur verið kennd við Háskóla Íslands allt frá stofnun hans árið 1911 og sama ár hófst kennsla við Alliance Française í Reykjavík. Franska er einnig kennd í íslenskum framhaldsskólum sem þriðja mál (sem er skyldugrein). Þrátt fyrir mikilvægi frönskunnar í íslensku skólakerfi skortir tilfinnanlega viðeigandi orðabækur og þá sér í lagi íslensk-franska orðabók þar sem slík orðabók hefur ekki komið út síðan *Íslenzk-frönsk orðabók* Boots var gefin út árið 1950.

Í þessari grein¹ verður hugað að því hvernig ný íslensk-frönsk orðabók gæti verið sem best úr garði gerð. Sjónum verður einkum beint að hlutverki íslensk-franskrar orðabókar með notendur og notagildi í huga. Í upphafi verður fjallað um þær tvímála orðabækur sem til eru fyrir tungumálin íslensku og frönsku og þá einkum *Íslenzk-franska orðabók*.² Reynt verður að lýsa því hvernig taka skal tillit til þarfa notenda við gerð nýrrar íslensk-franskrar orðabókar og hvaða áhrif það hefur á efnisval og framsetningu efnisins hvort slík orðabók

1 Ég þakka ónafngreindum ritrýnum tímaritsins *Orðs og tungu* fyrir gagnlegar athugasemdir og lagfæringar sem og ritstjóra þess, Ástu Svavarsdóttur, fyrir þarfar ábendingar og leiðsögn við ritun þessarar greinar. Jafnframt er Sigríði Stefánsdóttur þakkað fyrir yfirlestur.

2 Hér eftir verður vísað til þessarar orðabókar með skammstöfuninni *ÍFO*.

Orð og tunga 15 (2013), 73–93. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

er eingöngu ætluð íslenskum notendum eða hvort frönskumælandi notendur eru einnig hafðir í huga. Staldrað er við eitt helsta hlutverk tvímála orðabóka, það er að gefa upp jafnheiti á markmálinu fyrir hvert uppflettiorð viðfangsmálsins og hugað að því hvaða skilyrði þýðing uppflettiorðs þarf að uppfylla til að geta talist eiginlegt jafnheiti. Ólíkar notendaparfir verða skoðaðar með hliðsjón af fræðilegri umfjöllun um efnið og reynt að svara því hvernig hægt sé að taka tillit til beggja notendahópa, Íslendinga sem og frönskumælandi. Í því samhengi verða áhrif útgáfuformsins tekin til athugunar og úrlausnir sem raf-rænar orðabækur bjóða upp á kannaðar, einkum íslensk-norræna vef-orðabókin ISLEX og tvímála orðabókin *Le Grand Robert & Collins: dictionnaire anglais-français, français-anglais*.³ Dæmi verða tekin til að sýna upplýsingar sem ættu að koma fram í nýrri íslensk-franskri orðabók til þess að hún gefi sem besta mynd af notkun markmálsins innan ramma tvímála orðabókarlýsingar og geti þannig orðið betra tæki í tungumálanámi. Í viðauka er að finna nokkrar orðsgreinar í heild sinni. Þess ber að geta að ekki er um eiginlegt orðabókarverkefni að ræða heldur tillögur að orðsgreinum í mögulegri íslensk-franskri orðabók. Í lokaorðum verða niðurstöður dregnar saman.

2 Tvímála orðabækur milli íslensku og frönsku

Fyrstu orðabækurnar fyrir tungumálaparið íslensku og frönsku sem komu út voru bækur Páls Þorkelssonar *Íslenzk orðabók með frakkneskum þýðingum* frá árinu 1888⁴ og *Frönsk-íslenzk orðabók* frá 1914. Gérard Boots birti *Franskt-íslenzkt orðasafn* 1948 en síðar kom út eftir hann *Frönsk-íslenzk orðabók* (1953) með viðaukum eftir Þórhall Þorgilsson.⁵ Sú síðarnefnda er enn fánleg en notkun hennar hefur vikið fyrir nýrri fransk-íslenskri orðabók sem kom út árið 1995 og nefnist *Frönsk-íslensk orðabók*.⁶ Hún var gerð á grunni frönsku orðabókarinnar *Micro-Robert*

3 Hér eftir verður vísað til þessarar orðabókar með skammstöfuninni GRC. Enskur titill orðabókarinnar er *The Collins Robert French Dictionary: English-French, French-English*. Um er að ræða orðabók á CD-Rom disk frá 2008 en einnig er hægt að kaupa aðgang að orðabókinni á netinu.

4 Einungis 1. bindið (a-alblindur) kom út.

5 Á vefsíðu Stofnunar Árna Magnússonar í íslenskum fræðum er *Skrá um orðabækur og orðasöfn sem varða íslensku*, þar sem finna má yfirlit yfir velflestar orðabækur sem hafa verið gefnar út og varða íslensku.

6 Orðabókin var eingöngu gefin út á Íslandi og stoðmálið er íslenska sem bendir til þess að hún hafi einkum verið ætluð Íslendingum til að skilja frönsku.

(2. útgáfu 1988) sem er einmála orðabók ætluð frönskum nemendum og ekki hugsuð sérstaklega fyrir erlenda nemendur í frönsku. *Frönsk-íslensk orðabók*⁷ (1995) inniheldur um 35.000 orð en fyrir utan að þýða uppflettiorðin, notkunardæmi og orðasambönd var aukið við orðaförða úr *Micro-Robert*, einkum orðum úr viðskipta- og tæknimáli auk orða sem varða íslenska náttúru. Þessi orðabók er nú uppseld⁸ en er aðgengileg á vefbókasafninu *Snöru*. Í vefútgáfunni kemur hljóðritun uppflettiorða því miður ekki fram né heldur leturbreytingar, svo sem skáletrun notkunardæma og feitltrun fastra orðasambanda. *Frönsk-íslensk skólaorðabók* (1996) var unnin upp úr *FÍO* (1995) og er enn fáanleg. Einnig er til *Frönsk-íslensk, íslensk-frönsk orðabók: vasaorðabók* (Þór Stefánsson 1996)⁹ sem hefur að geyma um 15.000 flettur og er einkum ætluð byrjendum í frönskunámi en hana má flokka sem orðalista (e. *vocabulary*) þar sem engin dæmi eru gefin. Ætlunin var að ráðast í gerð íslensk-franskrar orðabókar í kjölfar þess að *FÍO* (1995) var samin en ekkert varð af þeim áformum.

Eina útgefna íslensk-franska orðabókin er því sem fyrr segir *ÍFO*. Enginn inngangur er að orðabókinni og því ekki vitað hvort hún var upphaflega ætluð bæði frönskumælandi og íslenskumælandi notendum eða einungis öðrum hópnnum. Orðabókin var aðeins gefin út á Íslandi og var því að öllum líkindum einkum ætluð Íslendingum. Hins vegar er stoðmálið franska og höfundur sleppir oftast að greina frá merkingarmun jafnheita á frönsku sem getur villt um fyrir íslenskum notendum. Þetta má sjá í (1) þar sem engin aðgreining er gerð á afturbeygðu sögninni *se perdre* og orðasambandinu *s'en aller en fumée* sem gefin eru upp sem þýðingar á orðasambandinu *fara í hundana*¹⁰ undir flettunni *hundur*:

(1) Flettan *hundur* í *ÍFO*:

hundur (-s, -ar) m.: chien m.; fara í hundana: se perdre, s'en aller en fumée.

Innskipan¹¹ orðabókarinnar er mjög einföld og ekki er stuðst við letur-

7 Hér eftir verður vísað til þessarar orðabókar með skammstöfuninni *FÍO*.

8 Árið 2000 kom út tölvuorðabók byggð á orðabókinni með 36.000 uppflettiorð (Áslaug Anna Þorvaldsdóttir 2010: 59).

9 Áður hafði vasaorðabókin *Frönsk-íslensk vasa-orðabók. Íslensk-frönsk vasa-orðabók* (1976) verið gefin út (sjá *Skrá um orðabækur og orðasöfn sem varða íslensku á vefsíðu Stofnunar Árna Magnússonar í íslenskum fræðum*).

10 Þar að auki eru þetta ekki heppilegar þýðingar á orðasambandinu *fara í hundana* og eru gerðar tillögur um aðrar þýðingar undir nafnorðsflettunni *hundur* í viðauka.

11 Með innskipan (e. *microstructure*) er átt við innri byggingu orðsgreina, það er

breytingar til þess að greina á milli ólíkra upplýsinga. Stundum eru engar málfræðiupplýsingar, hvorki um viðfangs- né markmál eins og lýsingarorðsflettan *brúnn* sýnir í (2):

(2) Flettan *brúnn* í *ÍFO*:

brúnn: brun; brúnn hestur: cheval noir.

Það er ljóst að *ÍFO* fullnægir ekki kröfum notenda á 21. öld þar sem orðaforði, bæði viðfangsmálsins (íslensku)¹² og markmálsins (frönsku), hefur breyst mikið á undanförunum 60 árum auk þess sem aðstæður til orðabókagerðar eru allt aðrar. Því verða íslenskir notendur sem vilja tjá sig á frönsku oft að fara í gengum þriðja tungumálið, svo sem ensku eða dönsku, og notast við til dæmis ensk-franska orðabók. Þetta fyrirkomulag torveldar íslenskum nemendum í frönsku aðgang að jafnheitum á frönsku þar sem þeir þurfa einnig að þekkja orðið sem þeir leita að á ensku. Ef þeir þekkja það ekki þurfa þeir fyrst að fletta upp í íslensk-enskri orðabók og því næst í ensk-franskri orðabók og eykur slíkur tvíverknaður hættuna á því að merkingin sem leitað var að týnist í þýðingu. Því er mikilvægt að orðabókanotendur hafi aðgang að tvímála orðabókum á móðurmáli sínu þegar þeir leitast við að tjá sig markvisst á erlendu máli.

3 Ný íslensk-frönsk orðabók

3.1 Hlutverk og notendur

Mikilvægt er að hafa í huga við gerð tvímála orðabókar hver ætla megi að verði virkasti notendahópur hennar og hvaða tilgangi hún eigi einkum að þjóna, hvort orðabókin eigi að gagnast notendum við að skilja erlent mál, talað eða skrifað, eða hvort hún eigi að nýtast þeim til að tjá sig á erlenda málinu. Í fyrra tilfellinu er talað um skilningsorðabækur eða L2→L1 orðabækur en í því síðara um málbeitingarorðabækur eða L1→L2 orðabækur (Sanders 2005:41, Marelló 1996:34). „L1“ stendur fyrir móðurmál notandans en „L2“ stendur fyrir erlenda tungumálið.¹³

hvernig lýsingu hvers flettiorðs fyrir sig er háttað en með meginskipan (e. *macro-structure*) er átt við ytri byggingu orðabókar (Jón Hilmar Jónsson 2005a: XIII).

12 Jafnframt er mikil áhersla á trúarlegan orðaforða í orðabókinni, t.d. má þar finna nafnorðsflettur *helgidagsbrot*, *kirkjumáldagi*, *kirkjusiðir*, *morguntíðir*, *passíusunnudagur*, *þakkarguðsþjónusta* o.s.frv.

13 „L1“ stendur fyrir „first language“ og „L2“ fyrir „second language“.

Líklegt er að notendur nýrrar íslensk-franskrar orðabókar yrðu aðallega Íslendingar. Gera má ráð fyrir að Íslendingar sem leggja stund á frönskunám séu fleiri en frönskumælandi notendur sem að einhverju marki hafa vald á íslensku þar sem franska hefur mikilvægari stöðu sem erlent tungumál í íslensku skólakerfi heldur en íslenska í Frakklandi. Þannig mætti setja málbeitingarhlutverk nýrrar íslensk-franskrar orðabókar í forgrunn og ætla að ný íslensk-frönsk orðabók nýttist fyrst og fremst íslenskumælandi notendum sem vilja tjá sig á frönsku munnlega eða skriflega. Yfirleitt er gengið út frá því að annað málið í tvímála orðabókum, annaðhvort viðfangsmálið (málið sem lýsingin beinist að) eða markmálið (málið sem þýtt er yfir á) sé móðurmál notandans (Svensén 2009:13) og þannig er gert ráð fyrir því að notendur orðabóka búi yfir ákveðinni færni í móðurmáli sínu sem nýtist þeim annaðhvort við að þýða af móðurmáli sínu eða yfir á móðurmál sitt (Kromann, Riiber og Rosbach 1991:2715; Sanders 2005:45). Sú er þó ekki raunin, þegar slíkar tvímála orðabækur vantar og íslenskir notendur verða að notast við til dæmis ensk-franskar eða dansk-þýskar orðabækur.

Samkvæmt þessari tvískiptingu eftir hlutverki og eftir því hvort móðurmál notandans er viðfangsmál eða markmál orðabókar má halda því fram að fyrir hvert tungumálapar þyrfti fjórar orðabækur. Þannig þyrfti eina íslensk-franska orðabók sem gagnaðist íslenskumælandi notendum sem málbeitingarorðabók (L1→L2) og aðra sem væri notuð sem skilningsorðabók af frönskumælandi notendum (L2→L1). Slík aðgreining er þó einungis fræðileg og óhagkvæm í reynd og yfirleitt skírskota tvímála orðabækur að einhverju leyti til ólíkra notkunarþarfa og hópa. Þannig er oft reynt að þjóna báðum notendahópum með orðabókum sem eru í senn skilningsorðabækur og málbeitingarorðabækur og er þá talað um tvíbeindar orðabækur (e. *bidirectional*) (Svensén 2009:18).¹⁴

ISLEX veforðabókin er dæmi um tvíbeinda orðabók en hún hefur íslensku sem viðfangsmál og norrænu málin dönsku, norskt bókmál, nýnorsku og sænsku sem markmál¹⁵ og beinist þannig í senn að norrænum notendum til skilnings á íslensku, t.d. þýðendum úr

14 Hugtakið *bifunctional* er einnig notað um orðabók sem gegnir tvöföldu hlutverki, t.d. tvímála orðabók frá tungumáli A til B sem er í senn ætlað að nýtast sem „aktív“ orðabók (til málbeitingar og við þýðingar af móðurmáli yfir á erlent tungumál) og sem „passív“ orðabók (við að skilja texta og við þýðingar af erlendu máli yfir á móðurmál) (NLO 1997: 85, Kromann 1990: 18).

15 Unnið er að því að bæta færeysku við sem markmáli.

íslensku yfir á eitthvert norrænu málanna (sem L2→L1 orðabók), og að Íslendingum sem vilja tjá sig á einhverju norrænu málanna (sem L1→L2 orðabók) (Aldís Sigurðardóttir *et al.* 2008:780).¹⁶ Tvímála orðabókin GRC beinist einnig að tveimur notendahópum, enskumælandi og frönskumælandi. Hún er jafnframt tvístefnuorðabók (e. *biscopal*) og sameinar þannig gerðirnar L1→L2 og L2→L1 en sú leið er oft farin með tvíbeindar orðabækur. Þannig gegnir hvor hluti fyrir sig tvöföldu hlutverki og í tilfalli GRC¹⁷ á ensk-franski hlutinn að hjálpa enskumælandi notendum að tjá sig á frönsku og þeim frönskumælandi að skilja ensku, og öfugt hvað varðar fransk-enska hlutann. GRC er ein víðamesta orðabókin fyrir tungumálaparið ensku og frönsku en hún inniheldur um 425.000 orð og orðasambönd (að báðum hlutum meðtöldum) og býður upp á fjölmörg dæmi um orðanotkun, bæði frjálsa og fasta.

Eins og fyrr segir er líklegast að íslensk-frönsk orðabók nýttist fyrst og fremst íslenskum frönskunemendum en þar sem töluverð gróska er í þýðingum á íslenskum bókmenntum yfir á frönsku og íslenska er kennd í nokkrum háskólum í Frakklandi verður skoðað hvernig íslensk-frönsk orðabók gæti einnig nýst frönskumælandi notendum og hvaða áhrif það hefði á val efnisatriða. Í því skyni verða fyrrnefndar orðabækur, ISLEX og GRC, hafðar til hliðsjónar í umfjöllun um nýja íslensk-franska orðabók.

3.2 Aðferðir og sjónarmið við val flettiorða og forsendur þess

Áætlaður notendahópur hefur áhrif á val flettiorða sem og á þær upplýsingategundir sem þurfa að koma fram í orðsgreinum orðabókar. Góð málbeitingarorðabók (L1→L2 orðabók) þarf að veita upplýsingar um hvernig á að nota jafnheitin á erlenda málinu og þurfa orðsgreinar slíkrar orðabókar því að innihalda notkunardæmi, merkingaradgreiningu og orðastæður. Þeir sem leitast við að tjá sig á erlendu máli þurfa frekar á almennum orðaforða að halda, orðaforða sem tengist daglegu lífi en síður á fornyrðum og staðbundnum orðaforða. Málbeitingarorðabækur ættu því að takmarkast við virkan og almennan orðaforða. Góð skilningsorðabók (L2→L1) ætti aftur á móti

¹⁶ ISLEX er byggð á gagnagrunni og inniheldur um 50.000 flettur og þýðingar á þeim (Aldís Sigurðardóttir *et al.* 2008: 780).

¹⁷ GRC er unnin af tveimur útgáfufyrirtækjum, Collins í Englandi og Les Dictionnaires Le Robert í Frakklandi.

að innihalda ríkulegri orðaforða en málbeitingarorðabækur (L1→L2), þ.á m. svæðisbundin orð og menningarbundin. Ástæðan er sú að málnotandi rekst mögulega á fjölbreyttari orðaforða í því sem hann les eða heyrir á erlenda málinu heldur en hann þarf á að halda þegar kemur að því að tjá sig á því. Það getur verið gagnlegt fyrir þýðendur úr íslensku yfir á frönsku að íslensk-frönsk orðabók innihaldi einnig eldri orðaforða sem finna má til dæmis í Íslendingasögunum (NLO 1997:84). Sem dæmi má nefna lýsingarorðið *frakkneskur* sem gæti verið gagnlegt fyrir frönskumælandi notendur en er í raun óþarft í íslensk-franskri orðabók eingöngu ætlaðri Íslendingum þar sem í íslensku máli í dag er lýsingarorðið *franskur* almennt notað.

Erfitt getur því verið að taka tillit til beggja notendahópa við val á orðaforða tvímála orðabókar og eins og Jón Hilmar Jónsson hefur bent á þá kallar það á margbreytilegan orðaforða ef taka á með inn í myndina þá notendur sem koma að viðfangsmálinu sem erlendu máli (2005b:23). Með því að hafa flettiorðaskrá íslensk-franskrar orðabókar sem ítarlegasta og takmarka hana ekki við algengan og almennan orðaforða eru meiri líkur á að orðabókin gagnist einnig frönskumælandi notendum og þýðendum úr íslensku yfir á frönsku. Þó má ætla að ný íslensk-frönsk orðabók geti takmarkast við orðaforða nútímamáls og því megi sleppa fornyrðum.

3.3 Jafnheiti og merkingaraðgreining

Meginhlutverk tvímála orðabóka er að gefa upp jafnheiti á markmálinu fyrir hvert uppflettiorð í viðfangsmálinu (Bogaards 2003:75). Hugtakið „jafnheiti“ er því lykilatriði við gerð tvímála orðabóka (Szende 1996:126). Tvímála orðabækur byggjast á því að stilla upp tveimur málkerfum samhliða og þannig verða orðabókahöfundar að finna jafnheiti á markmálinu sem samsvarar merkingu og notkun uppflettiorðsins á viðfangsmálinu. Jafnframt verða þeir að gæta þess að jafnheitin búi yfir sama merkingarblæ og tilheyri sama málsniði. Þetta fyrirkomulag gerir ráð fyrir því að alltaf sé hægt að þýða uppflettiorð yfir á annað tungumál. Hins vegar getur jafnheiti átt við þrengra merkingarsvið en uppflettiorðið á viðfangsmálinu eða þá að merking þess getur verið víðtækari (Jón Hilmar Jónsson 2005b:27).

Í raun er sjaldnast um fullkomið jafngildi (e. *full equivalence*) að ræða á milli tveggja tungumála (Svensén 2009:258). Til þess að hægt sé að líta á þýðingarorð sem eiginlegt jafnheiti þurfa hlutveruleikinn

eða táknmiðið (fr. *signifié*) að vera til staðar í menningu málnotenda og jafnframt þarf orðið sjálft eða táknmyndin (fr. *signifiant*) að vera til í tungumáli málnotendanna (Szende 1996:119). Sem dæmi má nefna jafnheitin *tölva* í íslensku og *ordinateur* í frönsku en í báðum tilfellum hafa orðin ótvíræða merkingu og vísa til sama hlutar. Þegar ekki finnst fullnægjandi jafnheiti verður orðabókahöfundur að útskýra merkingu flettiorðsins út frá samhengi (Szende 1996:119). Þetta er oft raunin þegar um menningarbundin orð er að ræða eins og til dæmis nafnorðið *skyr* sem á sér enga samsvörun í frönskum veruleika.

Ef ekkert eitt jafnheiti nær yfir margræða merkingu uppflettiorðs á viðfangsmálinu eru oft fleiri en eitt jafnheiti á markmálinu gefin upp sem þýðing. Í þeim tilvikum verður að aðgreina merkingu jafnheitanna til að gera notanda sem nálgast markmálið sem erlent mál kleift að átta sig á muninum og velja viðeigandi jafnheiti hverju sinni. Ekki nægir að aðgreina jafnheitin einungis með greinarmerkjum eins og kommu eða semikomu (Kromann, Riiber og Rosbach 1991:2724). Sem dæmi má nefna lýsingarorðið *brúnn* sem á sér tvö jafnheiti í frönsku: *brun* og *marron*, eftir því hverju er verið að lýsa. Lýsingarorðið *brun* er frekar notað þegar verið er að vísa til lífrænna hluta eins og moldar og hárs en óbeygjanlega lýsingarorðið *marron* er notað þegar vísað er til klæðnaðar og vefnaðar. Þó er vísað til augnlitar í Frakklandi¹⁸ með lýsingarorðinu *marron*, t.d. *il a des yeux marron(s)* 'hann er með brún augu' (Thibault 2007:476). Íslenskumælandi notendur þurfa að fá einhverja vísbendingu um það hvers vegna tvö jafnheiti eru gefin upp en frönskumælandi notendur geta valið á milli jafnheitanna hjálparlaust.

3.4 Upplýsingategundir

Áætlaðir notendur hafa ekki einungis áhrif á val flettiorða í tvímála orðabók heldur einnig á þær upplýsingategundir sem þurfa að koma fram í orðsgreinum. Ef erlenda málið er viðfangsmál orðabókarinnar (L2→L1 orðabók) þarf notandinn mun ítarlegri upplýsingar um er-

¹⁸ Þessi merkingaraðgreining lýsingarorðanna *brun* og *marron* á ekki við á öllum frönskumælandi svæðum, t.d. hefur notkun *brun* um augnit haldist í þeirri frönsku sem töluð er í Kanada og í Sviss: *il a des yeux bruns* (Thibault 2007: 476) en það er álitamál að hvaða marki málbeitingarorðabók á að taka tillit til þessa. Þetta er nýleg þróun og svo virðist sem sambandið *yeux marron(s)* sé að ryðja út sambandinu *yeux bruns* í Frakklandi.

lenda uppflettiorðið (t.d. kyn, óreglulega fleirtölu, kennimyndir sagna, andlag sagna) og notkun þess (til dæmis um málsnið, orðastæður) en um jafnheitin á móðurmáli sínu og notkun þeirra. Að sama skapi þarfnast notandi sem hefur viðfangsmálið að móðurmáli og nálgast markmálið sem erlent mál (L1→L2 orðabók) ítarlegri upplýsingar um jafnheitin og notkun þeirra heldur en um flettiorðið og notkun þess (Svensén 2009:15).

Ef íslensk-frönsk orðabók er einungis ætluð íslenskum notendum eru upplýsingar um setningagerð viðfangsmálsins (íslensku) óþarfar sem og upplýsingar um framburð og beygingu flettiorðanna. Aðalatriðið er að notendur fái nægilegar upplýsingar um jafnheitin á erlenda málinu og notkun þeirra því það er ekki nóg fyrir notandann að þekkja jafnheiti við tiltekið orð á markmálinu, hann verður líka að vita hvernig þau skuli notuð. Ef litið er aftur á lýsingarorðið *brúnn* og frönsk jafnheiti þess þá þarf íslenskur notandi til dæmis að vita að lýsingarorðið *marron* beygist ekki eftir kyni (ólíkt öðrum lýsingarorðum um liti) og fær yfirleitt ekki fleirtöluendinguna -s sem venjulega er bætt við nafnorð og lýsingarorð í fleirtölu. (Í viðauka má sjá orðsgreinina *brúnn* í heild sinni). Jafnframt má sýna með notkunardæmi að lýsingarorð yfir liti svo sem *brun* og *marron* koma á eftir nafnorðinu sem þau lýsa, t.d. *hún keypti brúna skó* 'elle a acheté des chaussures marrons(s)'.

3.5 Áhrif útgáfuformsins og breytilegt notendaviðmót

Lýsingin hér að framan hefur miðast við hefðbundnar prentaðar orðabækur. Í veforðabókum er orðabókartextinn hins vegar ekki fastur og óbreytanlegur og því má bæta jafnóðum við orðabókina. Jafnframt gefst meira rými fyrir skýringar og athugasemdir og þannig má ganga mun lengra en hingað til hefur verið gert hvað varðar upplýsingar um beygingu og setningarumhverfi flettiorðanna (Cartier 2000:139–140). Áhrif útgáfuformsins eru því töluverð en vandinn er engu að síður sá að því lengra sem gengið er í því að auka við upplýsingar um markmálið fyrir þá notendur sem ekki hafa það að móðurmáli, þeim mun meira er um óþarfar upplýsingar fyrir þann sem hefur markmálið að móðurmáli (Hannay 2003:150). Ómiklar upplýsingar á skjánum geta verið ruglandi fyrir notandann. Í ISLEX veforðabókinni er þetta leyst með því að sýna beygingarupplýsingar fyrir markmálin ef músar bendli er rennt yfir jafnheitið. Þar er jafnframt komið til móts við erlenda notendur með því að bjóða upp á tengingu við

Beygingarlýsingu íslensks nútímamáls (BÍN) þar sem notendur geta séð beygingu flettiorðsins í heild og þurfa ekki að leita annað í þeim tilgangi. Það væri kjörið að fara slíka leið í nýrri íslensk-franskri orðabók.

Þannig má hugsa sér nýja rafræna íslensk-franska orðabók þar sem notendur geta valið sér notendaviðmót eftir því hvort tungumál orðabókarinnar, viðfangsmálið (íslenska) eða markmálið (franska), er móðurmál þeirra (eða tungumálið sem þeir hafa betri tök á). Þetta má nýta til dæmis í sambandi við val á stoðmáli (e. *meta-language*), það er að segja því tungumáli sem notað er við skýringar og athugasemdir þar sem báðir notendahópar verða að geta skilið það sem þar er vísað til (*NLO* 1997:84). ISLEX veforðabókin býður notendum til dæmis upp á þann möguleika að velja notendaviðmót á viðeigandi tungumáli. Með sama móti má gera ráð fyrir að í rafrænni íslensk-franskri orðabók gætu Íslendingar valið íslensku sem stoðmál og frönskumælandi notendur frönsku. Rafrænar orðabækur geta einnig boðið upp á framburðardæmi (hljóddæmi) til viðbótar eða í stað hljóðritunar.

Hefð er fyrir því í orðabókagerð (bæði í prentuðum og rafrænum orðabókum) að sýna hljóðritun og/eða hljóddæmi eingöngu við uppflettiorðin og er það vissulega gagnlegt í L2→L1 orðabókum sem ætlaðar eru til viðtöku á erlendu máli. Í raun er ekki síður þörf fyrir framburðardæmi við jafnheitin í L1→L2 orðabókum sem ætlaðar eru til málbeitingar á erlenda málinu (Bogaards 2003:76; Svensén 2009:14). Gera má ráð fyrir að íslenskir notendur vilji vita um leið og þeir læra nýtt orð í frönsku hvernig orðið er borið fram, ekki síst þar sem íslenskir notendur hafa eins og fyrr segir ekki aðgang að framburði í *Fransk-íslenskri orðabók* á vefbókasafninu *Snöru*. Með því að bjóða upp á þann möguleika að hlusta á framburð jafnheita í nýrri íslensk-franskri veforðabók getur orðabókin orðið virkara tæki í tungumálanámi. Hins vegar er það plássfrekt og tímafrekt að setja inn hljóðritun á jafnheitum í hvert skipti sem þau koma fram í þýðingu á flettiorði og þar að auki getur það valdið vandkvæðum eins og þegar um fleiryrt jafnheiti er að ræða. Til dæmis getur sögnin *læra* ýmist verið þýdd með *étudier* eða *apprendre* en einnig með fleiryrta sagnorðssambandinu *faire des études de*.¹⁹

Svensén bendir á að þetta séu sterk rök fyrir því að gera tvístefnu-

¹⁹ Þar kemur líka til sögunnar svokölluð tenging í frönskum framburði (fr. *liaison*) þar sem *s* í lok óákveðna greinisins *des* raddast á undan á sérhljóða í upphafi næsta orðs (*études*) og er borinn fram [z]. Þessu þarf að gera grein fyrir í hljóðritun.

orðabækur, að minnsta kosti að samstæðar orðabækur um tiltekið tungumálapar byggist á sömu ritstjórnarstefnu og séu samdar á svipuðum tíma, til þess að orðabækurnar eða orðabókarhlutarnir styðji hvor við aðra (Svensén 2009:14). Í GRC er framburður til dæmis sýndur á hefðbundinn hátt við flettiorðin í hvorum hluta fyrir sig en notendur geta alltaf flett upp erlendu jafnheiti í hinum hlutanum og hlustað á framburð (í rafrænu útgáfunni) eða séð hljóðritun (í prentuðu gerðinni).

3.6 Tvíbeind orðabók – notkunardæmi og orðasambönd

Kromann hefur bent á að ef orðabók eigi að vera tvíbeind þurfi efnisatriði og upplýsingategundir orðsgreina fyrst og fremst að byggja á þörfum og forsendum málbeitingarorðabóka (L1→L2 orðabóka). Samkvæmt honum þurfa upplýsingar sem gefnar eru í orðsgreinunum þar af leiðandi að taka til merkingaraðgreiningar, málfræðiupplýsinga um markmálið, fastra orðasambanda, ófyrirsegjanlegra orðastæðna og menningarbundinna orða og gefa upp samsvaranir þeirra á markmálinu ásamt viðeigandi alfræðilegum upplýsingum á markmálinu (1990:24). Hér hefur þegar verið farið yfir mikilvægi merkingaraðgreiningar og málfræðiupplýsinga um markmálið. Mikilvægi menningarbundinna orða er augljóst þegar kemur að þýðingum af viðfangsmáli yfir á markmál en einnig getur verið gagnlegt út frá málbeitingarsjónarmiði að íslenskir notendur fái upplýsingar um menningarbundin heiti eins og til dæmis um það hvað hvaða menntun samsvari stúdentsprófi í Frakklandi eða hvaða franska stofnun sé sambærileg Mennta- og menningarmálaráðuneytinu.

Merking orðastæðna er sjaldnast yfirfærð en með ófyrirsegjanlegum orðastæðum á Kromann við að þótt orðastæður séu auðskiljanlegar erlendum notendum sé erfiðara fyrir þá að nota orðastæður þegar kemur að málbeitingu á erlendu tungumáli (Sanders 2005:45). Íslendingur sem vill finna samsvörun við íslensku orðastæðuna *að bursta tennurnar* á frönsku getur ekki vitað hvort hægt er að þýða hana beint eins og raunin er með frönsku orðastæðuna *se brosser les dents* eða hvort uppbyggingin sé allt önnur eins og er raunin með orðastæðuna *að læra heima* sem er á frönsku *faire les devoirs* 'gera heimavinnuna'. Oddný G. Sverrisdóttir (2009:164) bendir á það í grein sinni um eiginleika og flokka fastra orðasambanda að orðastæður á borð við *að bursta tennurnar* séu algengar í daglegu máli og þar af leiðandi mikilvægar þeim sem eru að læra erlent tungumál. Það er því full ástæða til að gera orðastæðum

betri skil í tvímála orðabókum en almennt hefur tíðkast eins og leitast er við að gera í ISLEX veforðabókinni en þar er rík áhersla á að sýna fjölbreytileg orðasambönd, fasta orðanotkun svo sem orðastæður og dæmi með þýðingum á markmálunum (Aldís Sigurðardóttir *et al.* 2008:780). Eins og Jón Hilmar Jónsson hefur bent á mætti auka við hefðbundna tvímála orðabókarlýsingu með orðastæðum á borð við *hundurinn geltir* og *hundurinn urrar* undir nafnorðsflettunni *hundur* (sjá orðsgrein í viðauka) þótt ekki sé hægt að gera ráð fyrir öðrum upplýsingum en þeim sem eiga við flettiorðið (2005b:30).

Orðastæður eru því sérstaklega mikilvægar í orðabók sem einkum er ætluð sem málbeitingarorðabók en öðru máli gegnir um orðasambönd. Ekki er ljóst af hverju Kromann telur að föst orðasambönd séu mikilvæg í orðabók sem eigi fyrst og fremst að miðast við þarfir notenda sem vilja tjá sig á erlendu máli. Ólíklegt er að notendur sem vilja koma hugsun sinni frá sér á öðru máli taki útgangspunkt í föstum orðasamböndum með yfirfærðri merkingu og þar að auki er óvíst að þau eigi sér beina samsvörun í markmálinu. Föst orðasambönd eru hins vegar mikilvæg þegar kemur að því að skilja og þýða texta og tilheyra því fremur skilnings- og þýðingarhlutverki tvímála orðabóka.

Til að taka tillit til bæði frönskumælandi og íslenskumælandi notenda í nýrri íslensk-franskri orðabók mætti hugsa sér að auka við lýsingu viðfangsmáls í tvímála orðabókum með því að veita aðgang að orðasamböndum í gegnum hugtakaheitalista, hvort heldur sem er á viðfangsmálinu eða markmálinu líkt og Jón Hilmar Jónsson hefur lagt til (2005b:36) og eins og gert er í *Stóru orðabókinni um íslenska málnotkun* (Jón Hilmar Jónsson 2005a). Anna Helga Hannesdóttir og Jón Hilmar Jónsson hafa einnig fjallað um hvernig auka mætti aðgang að orðasamböndum í íslensk-sænskri orðabók með slíkri hugtakaflokkun (2001). Með því að hafa hugtakaheitalistann einnig á erlenda tungumálinu er erlendum notendum jafnframt veittur beinni aðgangur að íslenskum orðasamböndum og orðastæðum. Sem dæmi hér má nefna ýmis orðasambönd sem venjan er að flokka undir nafnorðsflettuna *hundur* (sjá viðauka). Ef þessi orðasambönd, eins og til dæmis *að fara í hundana* eða *<eitthvað> fer í hund og kött*, eru tengd hugtakalista má finna þau með því að leita að hugtakinu „ógæfa“ og frönskumælandi notendur hefðu aðgang að þeim í gegnum franska hugtakið „malheur“.²⁰

20 Við val á orðasamböndum voru orðabækurnar ISLEX og *Stóra orðabókin um íslenska málnotkun* (2005) hafðar til hliðsjónar en einnig var tekið mið að

Tvímála orðabækur takmarkast stundum um of við það hlutverk sitt að gefa upp jafnheiti. Bogaards (2003:75) hefur bent á að vandamálið í L1→L2 orðabókum sé að notandinn sem vill tjá sig á markmálinu fái ekki nægilegar upplýsingar um einkenni og notkun jafnheitanna. Þetta einkennir sérstaklega einfaldari orðabækur sem ætlaðar eru byrjendum. Þar fá notendur oft afar knappar upplýsingar um markmálið þótt gera megi ráð fyrir að margir þeirra hafi litlar forsendur til að álykta um val á milli jafnheita. Notkunardæmi geta verið hjálpleg í þessu sambandi en þau gegna margþættu hlutverki, ekki aðeins því að staðfesta jafnheitin sem slík. Þau veita til að mynda mikilvægar upplýsingar um notkun jafnheitanna í ólíku samhengi. Þetta sýnir sig sérstaklega þegar kemur að notkun forsetninga. Til dæmis getur forsetningin *framan í* sem sýnd er í (3) svarað til forsetningarliðarins *sur le visage* í sambandinu „vera með <e-ð> framan í sér“, t.d. andlitsfarða, en aftur á móti er hún þýdd með *en face* t.d. í sambandinu „horfa framan í e-n“ (alla orðsgreinina má sjá í viðauka). Hér hefur forsetningin mismunandi merkingu eftir því hvaða falli hún stýrir. Þetta eru upplýsingar sem skipta máli fyrir frönskumælandi notendur en það má velta því fyrir sér hvort þetta ætti að koma fram í orðabók sem beinist einkum að Íslendingum:

(3)

hvað ertu með framan í þér?: qu'est-ce que tu as sur le visage?

ég þorði ekki að líta framan í hann: je n'ai pas osé le regarder en face

Notkunardæmi eru einnig notuð í því skyni að að greina á milli mismunandi merkinga uppflettiorðsins sem og í þeim tilgangi að veita hagkvæmar upplýsingar um svokallaðar samskiptaformúlur (Blanco 1996:103) sem eru föst orðasambönd notuð við ákveðnar aðstæður eins og *takk fyrir síðast* og *verði þér að góðu* (Oddný G. Sverrisdóttir 2009:157, 164). Oft eiga notkunardæmi að sýna upplýsingar sem koma fram annars staðar í orðsgreininni, til dæmis um beint andlag, og þannig hjálpa þau notandanum að skilja upplýsingarnar og að heimfæra þær þannig að hann geti notað uppflettiorðið á réttan hátt. Því er mikilvægt að tvímála orðabókarlýsing styðjist við fjölbreytt dæmi og var það sjónarmið haft að leiðarljósi við ritun orðsgreina fyrir nýja íslensk-franska orðabók. Sem dæmi má taka sagnorðið *læra* í (4) sem er ýmist þýtt sem *étudier* eða *apprendre* á frönsku eftir

einhverju leyti af GRC orðabókinni. Framsetning orðasambanda með afmörkun breytilegra liða innan oddklofa er einnig að fyrirmynd *Stóru orðabókarinnar um íslenska málnotkun* (2005).

Því hvert andlagið er. Sýnt er innan sviga dæmi um andlag, svo sem fag í háskóla eða tungumál, en notkunardæmi sýna þetta í samhengi (orðsgreinin er í heild sinni í viðauka):

(4) **læra**

so (+ þf.)

(+ grein í háskóla) **étudier / faire des études de**

hvað ertu að læra? qu'est-ce que tu fais comme études?

hún lærði hagfræði í Englandi elle a étudié l'économie en Angleterre

(+ tungumál, lexía) **apprendre**

hann lærir ensku í skólanum il apprend l'anglais à l'école

mig langar til að læra frönsku je voudrais apprendre le français

Þessar hugmyndir miða að því að víkka út tvímála orðabókarlýsingu og nýta þá vinnu sem lögð hefur verið í lýsingu hvers máls fyrir sig, í þessu tilfalli íslenskunnar. Form tvímála orðabóka er að vissu leyti takmarkandi (bæði á pappír og í rafrænu formi) þar sem lýsing markmálsins er bundin af lýsingu viðfangsmálsins og takmörk fyrir því hversu mikið af upplýsingum er hægt að veita um markmálið í tvímála orðabókum. Erfitt er að gera ráð fyrir meiri upplýsingum um markmálið heldur en um viðfangsmálið. Þó má leysa úr þessu upp að vissu marki til dæmis með tenglum og millivísunum í orðabókum sem innihalda tvo hluta L2→L1 og L1→L2. Hvað varðar nýja íslensk-franska orðabók má bæta úr þessu með fjölda notkunardæma og gefa þannig notandanum fleiri möguleika í málbeitingu sinni með dæmasetningum. Jafnframt má setja inn ramma með upplýsingum um menningarbundin orð og málfræðileg eða setningarfræðileg atriði markmálmegin og gera þannig mögulegt að lýsa markmálinu betur. Sem dæmi má nefna að í orðsgrein fyrir atviksorðið *heim* sem á sér jafnheitin *à la maison* og *chez soi* í frönsku væri tilvalið að vekja athygli íslenskra notenda (til dæmis með upplýsingum innan ramma) á því að sögnin *rentrer* er oftast notuð um það að „fara heim (til sín)” og atviksliðunum *à la maison* (‘heim’/‘heima’)²¹ og *chez soi* (‘til sín’) (5):

(5) <fara> *heim (til sín)*: *rentrer (à la maison/chez soi)*

21 Atviksorðið *à la maison* samsvarar bæði atviksorðunum *heim* og *heima* á íslensku.

4 Lokaorð

Í þessari grein hefur verið fjallað um hugmyndir um nýja íslensk-franska orðabók út frá hlutverki og ólíkum notendapörfum í því skyni að draga upp mynd af því hvernig slík orðabók gæti sem best komið til móts við íslenska frönskunemendur. Farið var stuttlega yfir þær orðabækur sem gefnar hafa verið út fyrir tungumálin íslensku og frönsku og þá einkum *Íslensk-franska orðabók (IFO)* sem er eina útgefna orðabókina með íslensku sem viðfangsmál og frönsku sem markmál. Nokkuð var fjallað um hefðbundna fræðilega skiptingu tvímála orðabóka eftir hlutverki þeirra, annars vegar í svonefndar skilningsorðabækur sem ætlaðar eru notendum við að skilja texta á erlendu tungumáli og við þýðingar af erlendu máli yfir á móðurmál og hins vegar í svonefndar málbeitingarorðabækur sem ætlaðar eru notendum við að tjá sig á erlendu tungumáli. Slík flokkun er þó einungis fræðileg og orðabækur gegna því oft tvöföldu hlutverki og er þá um að ræða tvíbeindar orðabækur. Erfitt getur verið að taka fullt tillit til þessara ólíku notendaparfa við gerð tvímála orðabókar, til dæmis við val flettiorða, þótt slíkt sé að jafnaði reynt.

Einn helsti ágalli margra tvímála orðabóka hefur verið sá hversu knappar upplýsingar þær veita um jafnheitin og aðgreiningu þeirra. Þar geta rafrænar orðabækur bætt úr með nánari merkingaraðgreiningu og fjölbreyttari notkunardæmum þar sem plássleysi er ekki lengur fyrirstaða. Þar að auki ættu orðastæður á borð við *að læra heima* og *að bursta tennurnar* að skipa stóran sess í tvímála orðabók sem hefur málbeitingarsjónarmið í fyrirrúmi. Slíkar upplýsingar eru mikilvægar þegar kemur að því að tjá sig á erlendu tungumáli. Með breyttri aðstöðu við orðabókagerð er einnig hægara um vik að taka tillit til ólíkra notendahópa til dæmis með breytanlegu notendaviðmóti og með því að auka við upplýsingarnar sem almennt eru gefnar í orðabókum og gefa ekki einungis upp jafnheiti á markmálinu þótt það verði auðvitað ætíð aðalhlutverk tvímála orðabóka. Líklegast er að ný íslensk-frönsk orðabók nýtist fyrst og fremst íslenskum frönskunemendum. Það sjónarmið var haft að leiðarljósi í dæmum um hvaða upplýsingar þyrftu að koma fram í íslensk-franskri orðabók til þess að íslensk-frönsk orðabók gæti orðið öflugt tæki fyrir notendur sem leitast við að tjá sig á markmálinu en jafnframt gagnleg þeim sem koma að viðfangsmálinu sem erlendu máli. Í viðauka eru sýndar fimm orðsgreinar eins og þær gætu litið út í nýrri íslensk-franskri orðabók.


Ljóst er að tilfinnanleg þörf er fyrir nýja íslensk-franska orðabók og því óskandi að ráðist verði í gerð slíkrar orðabókar í náninni framtíð.

Viðauki

Hér eru sýnd dæmi um nokkrar orðsgreinar eins og þær gætu litið út í nýrri íslensk-franskri orðabók sem gefin væri út í rafrænu formi. Gert er ráð fyrir því hér að valið sé íslenskt notendaviðmót (eins og rætt var í kafla 3.5 um breytilegt notendaviðmót rafrænna orðabóka) og þar af leiðandi eru allar upplýsingar um málfræði markmálsins (frönsku) einungis á íslensku. Jafnframt er upplýsingum um framburð og beygingarendingar íslensku uppflettiorðanna sleppt þar sem þær eru óþarfar fyrir íslenska notendur. Orðflokkur uppflettiorðsins er þó sýndur enda er það mikilvægt til þess að greina á milli orða eins og t.d. *vinna* (no.) og *vinna* (so.). Hljóðritun jafnheita á frönsku er sýnd og gert er ráð fyrir því að í framtíðar orðabókinni verði hægt að hlusta á framburð þeirra.²² Ekki er um eiginlegt orðabókarverkefni að ræða en vonast er til að þau sjónarmið sem hér koma fram um notendaviðmót og þarfir íslenskra notenda komi að gagni við gerð nýrrar íslensk-franskrar orðabókar í framtíðinni.

brúnn

lo

1 a) (litur) **brun**(-e) [brœ], [brýne] 

hún er með brúnt hár

elle a les cheveux bruns/elle est brune

b) (augnlitur, föt, skór, leður) **marron** [marɔ̃] 

hann er með brún augu

il a les yeux^{KK} marron(s)

hún keypti brúna skó

elle a acheté des chaussures^{KVK} marron(s)


marron beygist ekki í kyni og er oftast notað í eintölu

c) (sólbrúnn) **bronzé**(-e) [brɔ̃zɛ] 

þau voru brún eftir fríð

ils sont revenus bronzés de leurs vacances

2 (hestalitur)

bai brun [bɛbrœ] 

hryssan er brún

la jument est bai brun ^{óbeygjanlegt}

allir hestarnir eru brúnir


tous les chevaux sont bai brun ^{óbeygjanlegt}

bai brun er samsett lýsingarorð en *bai* eitt og sér merkir 'jarpur'

²² Þetta er táknað með hljóðmerki fyrir aftan hljóðritun innan hornklofa.

framan í

fs + þolfall/þágufall

sur le visage [syrləvizaʒ] *hann beindi vasaljósinu framan í mig**hvað ertu með framan í þér?***horfa/líta framan í e-n***ég þorði ekki að líta framan í hann***gretta sig framan í <e-n>***barnið gretti sig framan í mig*

il a dirigé la torche sur mon visage

qu'est-ce que tu as sur le visage?

regarder <qn> en face / dans les yeux



je n'ai pas osé le regarder en face

faire une grimace à <qn>

l'enfant m'a fait une grimace

heim

ao

[koma, fara] (til heimilis) **à la maison** [alamɛzɔ]  / **chez soi** [ʃeswa] (til lands) **dans son pays** [dāsɔpei] **<fara> heim (til sín)** = rentrer [rãtre]  (chez soi/à la maison)Í frönsku er sögnin **rentrer** oft notuð ein og sér um það að *faralkoma heim* og *chez soi* eða *à la maison* sleppt, t.d.: *er hún komin heim?* elle est déjà rentrée?**á leiðinni heim***ég hitti hana á leiðinni heim*

(með báti, flugvél)

bjóða <e-m> heim (til sín)*hann bauð okkur heim til sín***fylgja <e-m> heim***hún fylgdi mér heim*

en rentrant à la maison/chez soi

je l'ai rencontrée sur le chemin du retour

sur le trajet du retour

inviter <qn> chez soi

il nous a invités chez lui

raccompagner <qn> à la maison

elle m'a raccompagné à la maison

hundur

no kk

1 a) (dýr) **chien**^{KK} [ʃjɛ] *fara út með hundinn**varið ykkur á hundinum!**hundurinn dillar rófunni**hundurinn flaðrar upp um <mig>**hundurinn geltir**hundurinn gjammar**hundurinn glefsar í <mig>**hundurinn spangólar**hundurinn urrar*

promener/sortir son chien

attention, chien méchant !

le chien remue la queue

le chien <me> fait fête


le chien aboie

le chien jappe

le chien <me> mordille

le chien hurle

le chien grogne

b) (ópokki, illmenni) **canaille**^{KVK} [kanaj] *hundurinn þinn!*

sale canaille !

2 ORÐASAMBÖND**<eitthvað> fer í hund og kött***fara í hundana**rífast eins og hundur og köttur**vera eins og hundur og köttur**það er ekki/varla hundi út sigandi*

<quelque chose> va à vau-l'eau

tourner mal

s'entendre comme chien et chat

être/vivre comme chien et chat

il fait un temps à ne pas mettre un chien

dehors


það er hundur í <honum>
þar liggur hundurinn grafinn
 → **rauðir hundar**²³: la rubéole

<il est> d'une humeur de chien
 c'est là le cœur du problème

læra

so + þolfall

(+ grein í háskóla)

étudier [etydjɛ] , **faire des études de** [fɛrdezetyddə] 


hvað ertu að læra?

qu'est-ce que tu fais comme études?

hún lærði hagfræði í Englandi

elle a étudié l'économie en Angleterre

(+ tungumál, lexía)

apprendre [apɾɑ̃dɾ] 

hann lærir ensku í skólanum

il apprend l'anglais à l'école

mig langar til að læra frönsku

je voudrais apprendre le français

finnast gaman að læra

avoir le goût de l'étude

vera fljót(ur) að læra

apprendre vite/facilement

læra (*sína*) **lexíu** (óeiginl. merking)

comprendre la leçon (óeiginl. merking)

hann hefur lært sína lexíu

il a compris la leçon

læra að <lesa>

apprendre à <lire>

læra á bíl

apprendre à conduire

læra á <píanó>

apprendre <le piano>

læra (<e-ð>) **utanbókar / utan að**

apprendre (<qch.>) par cœur

læra fyrir / undir próf

réviser

ég þarf að læra fyrir prófið á morgun

il faut que je révise pour l'examen de
demain

læra heima

faire ses devoirs

læra af reynslunni

apprendre de/par l'expérience

læra af mistökunum

tirer la leçon de ses erreurs

læra um <e-ð>

étudier qch.

Heimildir

Aldís Sigurðardóttir, Anna Hannesdóttir, Halldóra Jónsdóttir, Hákan Jansson, Lars Trap-Jensen og Þórdís Úlfarsdóttir. 2008. ISLEX – an Icelandic-Scandinavian Multilingual Online Dictionary. Í: E. Bernal og J. DeCesaris (ritstj.). *Proceedings of the XIII EURALEX International Congress* (Barcelona, 15–19 July 2008). http://www.euralex.org/elx_proceedings/Euralex2008/. (03.09.2012).

Anna Helga Hannesdóttir og Jón Hilmar Jónsson. 2001. Að hafa í sig og á. Ísländsk fraseologi í ett isländsk-svenskt perspektiv. *LexicoNordica* 8:67–91. Áslaug Anna Þorvaldsdóttir. 2010. „Orð er á Íslandi til“ *Um tvímála orðabækur og þýðingar*. Óprentuð ritgerð til M.A.-prófs í þýðingfræði við Háskóla Íslands.

BÍN = *Beygingarlýsing íslensks nútímamáls*. Á vefsíðu Stofnunar Árna Magnússonar í íslenskum fræðum: <http://bin.arnastofnun.is/>. (03.09.2012).

23 Hér ert gert ráð fyrir vísun í flettuna *rauðir hundar*.

- Blanco, Xavier. 1996. Élaboration et réutilisation des exemples dans la lexicographie bilingue. Í: H. Béjoint og P. Thoiron (ritstj.). *Les Dictionnaires bilingues*, bls. 103–110. París: Duculot.
- Bogaards, Paul. 2003. La Répartition des données dans les dictionnaires bilingues. Í: T. Szende (ritstj.). *Les Écarts culturels dans les dictionnaires bilingues. Actes des 3^{ème} journée sur la lexicographie bilingue*, bls. 75–86. París: Honoré Champion.
- Bogaards, Paul. 2006. Produire en L2 au moyen d'un dictionnaire bilingue. Í: T. Szende (ritstj.) *Le français dans les dictionnaires bilingues*, bls. 23–34. París: Honoré Champion.
- Boots, Gérard. 1948. *Franskt-íslenskt orðasafn*. Reykjavík: Ísafoldarprentsmiðja.
- Boots, Gérard. 1953. *Frönsk-íslensk orðabók*. Með viðaukum eftir Þórhall Þor­gilsson. Reykjavík: Ísafoldarprentsmiðja.
- Boots, Gérard. 1950. *Íslensk-frönsk orðabók*. (1. útgáfa) Reykjavík: Ísafoldarprentsmiðja.
- Cartier, Emmanuel. 2000. Éléments pour une modélisation des dictionnaires électroniques. Í: Th. Szende (ritstj.). *Dictionnaires bilingues. Méthodes et contenus*, bls. 135–152. París: Honoré Champion.
- FÍO = *Frönsk-íslensk orðabók*. 1995. (1. útgáfa) Ritstjóri: Þór Stefánsson. Reykja­vík: Örn og Örlygur; París: Dictionnaires Le Robert.
- Frönsk-íslensk vasa-orðabók*. *Íslensk-frönsk vasa-orðabók*. 1976. Ritstjórar: Elín­borg Stefánsdóttir og Gérard Chinotti. Reykjavík: Orðabókaútgáfan.
- Frönsk-íslensk skólaorðabók*. 1999. (1. útgáfa) Ritstjóri: Hrafnhildur Guðmunds­dóttir. Reykjavík: Mál og menning.
- GRC = *Le Grand Robert & Collins: dictionnaire anglais-français, français-anglais*. 2008 (rafræn útgáfa á geisladiski). París: Dictionnaires Le Robert.
- Hannay, Mike. 2003. Types of bilingual dictionaries. Í: Piet van Sterkenburg (ritstj.). *A Practical Guide to Lexicography*, bls. 145–153. Amsterdam, Phila­delphia: John Benjamins.
- ÍFO = Gérard Boots. 1950. *Íslensk-frönsk orðabók*. (1. útgáfa) Reykjavík: Ísa­foldarprentsmiðja.
- Íslensk orðabók*. 2002. (3. útg. aukin og endurbætt) Ritstjóri: Mörður Árnason. Reykjavík: Edda.
- ISLEX. Ritstjóri: Þórdís Úlfarsdóttir. <http://www.islex.hi.is/>. (03.09.2012).
- Jón Hilmar Jónsson. 2001. Staða orðasambanda í orðabókarlýsingu. *Orð og tunga* 5:61–86.
- Jón Hilmar Jónsson. 2005a. *Stóra orðabókin um íslenska málnotkun*. Reykjavík: JPV-útgáfa.
- Jón Hilmar Jónsson. 2005b. Aðgangur og efnisskipan í íslensk-erlendum orðabókum – vandi og valkostir. *Orð og tunga* 7:21–40.
- Kromann, Hans-Peder. 1990. Selection and presentation of translational equivalents in monofunctional and bifunctional dictionaries. *Cahiers de Lexicologie* vol. 56–57. Actes du colloque franco-danois de lexicographie. Copenhague, 19 et 20 septembre 1988, bls.17–26.

- Kromann, Hans-Peder; Theis Riiber og Poul Rosbach. 1991. Principles of Bilingual Lexicography. Í: F. J. Hausmann *et al.* (ritstj.). *Wörterbücher. Ein internationales Handbuch zur Lexikographie* 3. bindi, bls. 2711–2728. Berlin, New York: Walter de Gruyter.
- Marello, Carla. 1996. Les Différents types de dictionnaires bilingues. Í: H. Béjoint og P. Thoiron (ritstj.). *Les Dictionnaires bilingues*, bls. 31–52. París: Duculot.
- NLO = *Nordisk leksikografisk ordbok*. 1997. Ritstjórar: Henning Bergenholtz *et al.* Oslo: Universitetsforlaget.
- Oddný G. Sverrisdóttir. 2009. Orð til taks. Af eiginleikum og flokkum fastra orðasambanda. Í: Magnús Sigurðsson og Rebekka Þráinsdóttir (ritstj.). *Milli mála. Ársrit Stofnunar Vigdísar Finnbogadóttur í erlendum tungumálum* 1, bls. 149–171. Reykjavík: Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum.
- Páll Þorkelsson. 1888. *Íslensk orðabók með frakkneskum þýðingum*. Bindi 1,1 (a-alblindur). Reykjavík: Páll Þorkelsson.
- Páll Þorkelsson. 1914. *Frönsk-íslensk orðabók*. Reykjavík: Guðm. Kr. Guðmundsson.
- Sanders, Christopher. 2005. Bilingual Dictionaries of Icelandic: Types of Users and their Different Needs – a Discussion. *Orð og tunga* 7: 41–57.
- Skrá um orðabækur og orðasöfn sem varða íslensku*. Á vefsíðu Stofnunar Árna Magnússonar í íslenskum fræðum: http://www.arnastofnun.is/page/arnastofnun_ord_ordabokaskra (03.09.2012).
- Snara*. <http://snara.is/8/>. (03.09.2012).
- Svensén, Bo. 2009. *A Handbook of Lexicography. The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.
- Szende, Thomas. 1996. Problèmes d'équivalence dans les dictionnaires bilingues. Í: H. Béjoint og P. Thoiron (ritstj.). *Les Dictionnaires bilingues*, bls. 111–126. París: Duculot.
- Thibault, André. 2007. Banques de données textuelles, régionalismes de fréquence et régionalismes négatifs. Í: David Trotter (ritstj.). *Actes du XXIV^e Congrès International de Linguistique et de Philologie Romanes, Aberystwyth 2004*, 1. bindi, bls. 467–480. Tübingen: Max Niemeyer.
- Þór Stefánsson. 1996. *Frönsk-íslensk, íslensk-frönsk orðabók: vasaorðabók. Français-islandais, islandais-français dictionnaire de poche*. Reykjavík: Orðabókaútgáfan.
- Þórdís Úlfarsdóttir. 2006. ISLEX – íslensk-norræn veforðabók. *Orð og tunga* 8:147–148.

Lykilorð

tví mála orðabækur, íslensk-frönsk orðabók, jafnheiti, orðastæður, notkunardæmi.

Keywords

bilingual lexicography, Icelandic-French dictionary, equivalents, collocations, example sentences.

Abstract

This article describes the role of an Icelandic-French dictionary focussing on the expected users of such a dictionary. The existing dictionaries for the language pair, Icelandic and French, are briefly presented, in particular the only published dictionary with Icelandic as the source language and French as the target language, *Íslensk-frönsk orðabók* (Boots 1950). The role of dictionaries according to their uses either for production or reception of a foreign language is examined. The characteristics of bilingual dictionaries are discussed, especially the conditions that equivalents need to fulfil as well as the importance of meaning discrimination when more than one equivalent is proposed for a given headword. The selection of headwords and information types in bidirectional dictionaries are considered and the microstructure of the entries. The possibilities and solutions that electronic dictionaries offer are taken into account, for example regarding the user interface. In this context the Icelandic-Scandinavian dictionary ISLEX is considered as well as the French-English, English-French dictionary *Le Grand Robert & Collins*. Examples are shown for an Icelandic-French Dictionary (and elaborated entries are proposed in the appendix), where the focus is on inclusion of more examples and collocations in order to illustrate the use of equivalents in the target language.

Rósa Elín Davíðsdóttir
doktorsnemi í orðabókafræðum
Université Paris-Sorbonne (Paris IV) /
Deild erlendra tungumála, bókmennta og málvísinda, Háskóla Íslands
red1@hi.is



Guðrún Kvaran

Orðabókarstörf Konráðs Gíslasonar

1 Inngangur

Flestum kemur líklega fyrst í hug málfræði, tímaritið *Fjölirnir* og Jónas Hallgrímsson þegar þeir heyra nafn Konráðs Gíslasonar. Færri minnst hans sem orðabókarmanns og kom hann þó að tveimur útgefnum bókum og hafði fleiri um tíma í undirbúningi. Í greininni verður fjallað um orðabókarmanninn Konráð Gíslason. Fyrst verður dregið saman það sem vitað er um aðkomu hans að orðabókinni *An Icelandic-English Dictionary* sem gefin var út í heftum á árunum 1869–1874. Af orðum hans sjálfs að dæma í bréfi frá 1872 til danska kirkju- og menntamálaráðuneytisins (sjá 2. kafla) mun hann hafa skipulagt þá bók og haft umsjón með henni lengi framan af verkinu.

Í þriðja kafla verður rætt um *Danska orðabók með íslenskum þýðingum* sem prentuð var í Kaupmannahöfn 1851 og er eina orðabókin sem Konráði tókst að fylgja eftir til loka. Þess vegna verður henni gefið meira rými en íslensk-ensku orðabókinni. Sýnishorn verður skoðað og borið að þeim orðabókum sem Konráð nefnir í formála að hann hafi notað við verkið. Reynt verður að gera grein fyrir vinnubrögðum hans þótt önnur lýsing hafi ekki fundist en sú sem fram kemur í nokkuð ítarlegum formála fyrir verkinu. Hann nefnir bókina t.d. hvorki í bréfi sem hann skrifaði konungi 1847 (sjá 2. kafla) né í því sem vísað var til hér að ofan.

Síðar í greininni verður stuttlega dregið saman það sem vitað er um önnur orðabókarverk sem Konráð hafði í huga eða vann við um lengri eða skemmri tíma. Til er lýsing hans sjálfs í bréfinu til kirkju- og menntamálaráðuneytisins frá 1872 á fornorrænni-danskri orðabók

Orð og tunga 15 (2013), 95–119. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

og er áhugavert að kynnast hugmyndum hans um vinnu við slíka orðabók þótt handritið virðist hafa glatast.

2 Íslensk-ensk orðabók

Fyrsta orðabókarverkið sem Konráð tók þátt í að semja var íslensk-ensk orðabók. Hann hóf verkið snemma árs 1840 og var ráðinn til þess af breska auðmanninum Richard Cleasby. Með honum vann Brynjólfur Pétursson fyrst í stað en síðar komu fleiri að verkinu. Til er bréf frá Konráði til Danakonungs frá 13. desember 1847¹ þar sem fram kemur að hann hafi unnið samfellt við orðabókina frá maí 1840 til september 1847. Í bréfi sem Hallgrímur Scheving skrifaði Konráði 4. ágúst 1845 er þó ljóst að Konráð hætti um tíma að vinna fyrir Cleasby á þessum árum, Hallgrími til sárrar mæðu. Í bréfinu stendur m.a. (Finnbogi Guðmundsson 1970:184):

Mikið sé eg eftir því, að þér yfirgáfuð Cleasby, meðan hann þurfti yðar við, því fyrir það sama er hætt við, að talsvert fleiri misfellur verði á verki hans en annars hefði orðið. En að maðurinn hafi þurft yðar við, þó hann ef til vill af enskri stórmennsku ekki hafi viljað láta á því bera, hefir hann sýnt með því að vera sér úti um nýja tilhjálparmenn.

Í öðru bréfi frá Hallgrími frá 3. mars 1847 sést að Konráð er aftur kominn til starfa við orðabókina og spyrst Hallgrímur fréttu af gangi verksins (Finnbogi Guðmundsson 1970:186).

Cleasby lést 1847. Í formála að *Oldnordisk ordbog* eftir Eirík Jónsson (1863:XVIII) kemur fram að söfnun til orðabókarinnar hafi þá að mestu verið lokið og þýðingar yfir á ensku þegar hafnar. Þeim hafi verið haldið áfram og komu að þeim, auk Konráðs, nokkrir Íslendingar þar á meðal Eiríkur sjálfur.

Í mars 1854 taldi Konráð verkið langt komið og sendi prentað sýnishorn til þýska málfræðingsins og orðabókarmannsins Jakobs Grimm til þess að fá álit hans. Grimm hafði sjálfur unnið um nokkurt skeið að þýskri orðabók, *Deutsches Wörterbuch*, og kom fyrsta bindi hennar einmitt út þetta sama ár. Í bréfi frá 5. maí sama ár fór Grimm lofsamlegum orðum um sýnishornið sem Konráð sendi, lýsti yfir ánægju sinni með að Konráði skyldi hafa verið falið að fylgja verkinu eftir og

¹ Afrit af bréfinu er varðveitt á Stofnun Árna Magnússonar í Kaupmannahöfn (Árnasafn, KG 31 a).

spyr hvort eitt bindi muni virkilega nægja fyrir svo umfangsmikið verk (Briefwechsel 1885:309–311).

Sama ár og Konráð sendi sýnishornið til Jakobs Grimm kröfðust erfingjar Cleasbys að fá handritið sent til Englands. Þeir gáfu það Oxfordháskóla og kom það síðar í hlut George Webbe Dasents og Guðbrands Vigfússonar að búa verkið til útgáfu. Konráðs var að engu getið í fyrstu útgáfu orðabókarinnar sem kom út í heftum í Oxford 1869–1874. Hlutur Konráðs og annarra Íslendinga, sem unnið höfðu að bókinni, var réttur þegar Sir William Craigie kom að annarri útgáfu (Cleasby-Vigfússon 1957:vi).

Björn M. Ólsen skrifaði minningargrein um Konráð í *Tímarit Hins íslenska bókmenntafélags* 1891 og fer þar hörðum orðum um Cleasby og samning hans við Konráð (68–71). Björn þekkti Konráð vel og var samtímis honum í Kaupmannahöfn á árunum 1872–1877. Er því ekki ólíklegt að hann hafi haft í huga frásögn Konráðs sjálfs þegar hann skrifaði (68–69):

Árið 1839 í nóvembermánuði kom til Kaupmannahafnar enskur maður að nafni Richard Cleasby. Hann langaði til að læra íslensku og fjekk kenslu hjá Konráði Gíslasini. Cleasby var auðmaður, enn Konráð fjelítill, og komu þeir sjer nokkru síðar (í aprílmánuði 1840) saman um það, að Konráð skildi fara að safna til forníslenskrar orðabókar, sem svo skildi koma út á kostnað Cleasbys og með nafni hans, enn Cleasby skildi borga Konráði 40 ríkisdali á mánuði hverjum fyrir starf hans. Það var einn af þessum vanalegu samningum milli auðmannsins, sem ekkert veit nje getur, og fátæklingsins, sem hefur góða hæfilegleika; auðmaðurinn lætur peningana, sem hann hefur, fyrir það starf, sem hann vantar vit og þekkingu til að leisa af hendi, og eignar sjer svo að sjálfsögðu allan sóma af starfinu, sem hann hefur keipt, og oft hafa fátækir Íslendingar fir og síðar undirskrifað slíka samninga.

Síðan rekur hann að Konráð hafi unnið fyrir Cleasby í sjö ár og fyrir erfingjana í önnur sjö en þá hafi hann þurft að senda safnið hálfkarað til Englands. Eftir það rekur Björn aðkomu Guðbrands Vigfússonar að verkinu og segir (70):

Og það er að minsta kosti óhætt að fullirða, að Cleasby á sjálfur engan staf í orðabókinni, og að sýnishorn þau, sem prentuð eru í formála orðabókarinnar og eignuð Cleasby, eru öll frumsamin af Konráði, þó að þau sjeu prentuð eftir afskriftum með hendi Cleasbys. [...] Enn enginn getur annað sagt, enn að hann hafi fengið fult endurgjald fyrir það, sem hann ljet út, í því, sem Konráð vann fyrir hann, og

aldrei mundi Cleasby, ef hann hefði lifað, hafa gert sig sekan í eins hróplegu vanþakklæti og þeir gerðu, sem rituðu formálann fyrir orðabókinni eftir lát Cleasbys, því að þeir dróttu því að Konráði og öðrum Íslendingum, sem að orðabókinni unnu, að þeir hafi haft Cleasby fyrir fjeþúfu og lítið gert, og reina þannig eftir fremsta megni að svívirða þann mann, sem mest hafði unnið að undirbúningi orðabókarinnar.

Orð Björns M. Ólsens eru hörð og óvægin en til er umsögn Konráðs sjálfs í bréfi dagsettu 26. september 1872 „Til det høje Ministerium for Kirke- og Undervisningsvæsenene.“² Í bréfinu afsakar hann dráttinn á forníslensk-danskri orðabók (sjá 4. kafla) með því m.a. að telja upp þær orðabækur sem út hafi komið frá 1860. Um íslensk-ensku orðabókina skrifaði hann:

Richard Cleasby's „Icelandic-English Dictionary“ (A-R. Oxford 1869–1871) er fra først af udarbejdet efter min Plan og under min Opsigt, ligesom jeg ogsaa selv har samlet en stor Del af det i samme indeholdte Materiale. Men denne Ordbog „enlarged and completed by Gudbrand Vigfusson“ frembyder efter min Mening i sin nuværende Form (trods den Belæring, der endnu kan hentes) i det hele taget et sørgeligt Exempel paa Tilbagegang i kritisk Henseende. Forøgelsen og Fuldendelsen bestaar væsentligst i Tilføjning af uvedkommende nysislandske Ord og Talemaader, samt nogle af citatløse Egennavne, og Opfatningen synes ikke at vidne om klar Indsigt i Oldtidens Ord og Tanke.

Af þessum orðum er ljóst að Konráð var síður en svo ánægður með það verk sem hann í upphafi skipulagði og hafði umsjón með. Hann virðist hafa haft ánægju af því að vinna að íslensk-ensku orðabókinni á meðan sú vinna stóð og sést það m.a. af því sem hann skrifaði í bréfi til föður síns 26. september 1850 (Bréf 1984:150):

Jeg hef ella mikið að gjöra. Dönsk orðabók, æði stór, handa Íslendingum, á (ef g. l.) að koma út frá mjer, í endalok næsta mánaðar (októbers) eða í nóvember. Þegar það er búið, fer jeg (ef g. l.) að halda áfram hinni íslenzku orðabók handa enskum mönnum, sem Cleasby heitinn (enskur maður) hafði byrjað; og þykir mjer meira varið í það starf, en hitt (með dönsku orðabókina).

Edvard Collin etatsráð (1808–1886) og Konráð skrifuðust á um orðabók Cleasbys á árunum 1872 til 1884. Í bréfi dagsettu 27. nóvember 1872 afsakar Konráð drátt á svari en gerir síðan ítarlegar athugasemdir við

2 Vitnað er í ljósrít af bréfi sem varðveitt er í Konunglega bókasafninu í Kaupmannahöfn.

einstakar greinar í orðabók Cleasbys frá blaðsíðu 249 til 493. Í öðru bréfi frá 30. nóvember sama ár eru viðbætur frá blaðsíðu 352 til 393. Ekki er ólíklegt að Konráð hafi verið búinn að skrifa áður um fyrri hluta orðabókarinnar en bréf um það fundust ekki. Svör Konráðs eru fræðilega unnin, ekki örlar á gremju, en ljóst er að hann hafði ýmislegt út á einstök atriði að setja.

Það verður ef til vill seint hægt að sýna fram á hvaða hlut Konráð, og reyndar aðrir Íslendingar líka, átti í íslensk-ensku orðabókinni, en ef marka má það sem dregið var fram hér að framan er hann allmikill. Óumdeilanlegur er hlutur hans í þeirri orðabók sem rætt verður um í næsta kafla.

3 *Dönsk orðabók með íslenskum þýðingum*

Nánast á sama tíma og Konráð tók að vinna að íslensk-enskri forn-málsorðabók fyrir Cleasby hófst hann handa við að semja *Danska orðabók með íslenskum þýðingum*. Upphaflega átti orðabókin að verða samvinnuverkefni þeirra Konráðs og Jóhanns Halldórssonar sem dvaldist í Kaupmannahöfn frá 1834 til dauðadags 1844 (Páll Eggert Ólason III 1950:24–25). Í formála kemur fram að þeir skiptu verkinu þannig að Jóhann átti að sjá um stafkaflann *m-ø* og hafði lokið við *m* og mestan hluta af *n* þegar hann féll óvænt frá. Eftir það sinnti Konráð verkinu einn og segir í formála (1851:III):

Eftir lát hans breytti jeg nokkuð öllu sniði bókarinnar (eins á þeim hlutanum, sem jeg hafði unnið sjálfur), en hagnýtti það sem hann var búinn með, og vitnaði í það þar sem mjer þótti fallið.

Tilvísanir til Jóhanns eru merktar með skammstöfuninni Jóh. Konráð lætur þess einnig getið að Jóhann hafi verið hvatamaður þess að sótt var um styrk til verksins. Bókin kom út í Kaupmannahöfn 1851. Hún er 596 blaðsíður í stóru broti, þrídálka.

Formáli bókarinnar er mjög rækilegur og gerir Konráð þar grein fyrir þeim orðabókum sem hann notaði. Dönskuna segist hann að mestu hafa tekið úr *Dansk Ordbog* frá 1833 eftir Christian Molbech. Samanburður við hana sýnir að Konráð hefur fylgt henni nokkuð ítarlega. Sumt segist hann hafa tekið úr *Dansk Ordbog udgiven under Videnskabernes Selskabs Bestyrelse* í sex bindum frá 1793–1848 en útgáfu þess verks lauk ekki fyrr en með áttunda bindi 1905 þannig að Konráð hafði ekki aðgang að síðasta hluta stafrófsins. Þá nefnir hann *Fremmed-*

Ordbog eftir Ludwig Meyer frá 1844 sem er önnur útgáfa þeirrar bókar en þangað sótti hann tökuorð í dönsku.

3.1 Samanburður við orðabók Molbechs

Vinnubrögð Konráðs má best sjá af samanburði við þær heimildir sem hann tilgreinir. Í formálanum fyrir orðabók sinni lýsir hann nokkuð nákvæmlega hvernig hann fór að við framsetningu greina, hverju hann breytti frá orðabók Molbechs (t.d. stafsetningu), hverju hann hélt eftir og hvernig flettugreinar eru settar fram, t.d. hvernig greinarmerki eru notuð, og hvað stendur innan oddklofa eða sviga. Hann nefnir í þessu sambandi ekki orðabók Vísindafélagsins.

Í skýringum notar hann víða tilvísanir til fornrita (sjá t.d. í töflu 1 undir *Deeltagelse* og *Deger*). Um þessa ákvörðun sína skrifar Konráð (IV):

Einhver helzti kostur á bók þessari væntir mig að það sje, er jeg hef vísað í fornþækur eða tekið dæmi úr þeim, þó mjer þætti ekki þörf á að þræða rithátt þeirra. – „(fornt)“, við íslenzkt orð, merkir ekki, að þau sjeu úrelt, heldur að þau finnist í fornum bókum.

Á sama tíma og hann er að vinna við dönsku orðabókina er hann að störfum við *Icelandic-English Dictionary* (sbr. kafla 2) og hefur því væntanlega haft dæmi á hraðbergi sem að gagni komu að hans mati.

Rúmsins vegna valdi ég aðeins af handahófi fyrsta dálk á síðu 80 í bók Konráðs, frá *Deeltagelse* til *deiet*, og bar hann saman við orðabók Molbechs og orðabók Vísindafélagsins frá *Deeltagelse* til *Degne* (sbr. töflu 1) og eftir það aðeins við bók Molbechs (sbr. töflu 2) þar sem í ljós kom að bein tengsl virtust óveruleg við orðabók Vísindafélagsins en aftur á móti talsverð við bók Molbechs. Flettiorðin í bók Molbechs á síðum 160–161 eru nær öll hin sömu og hjá Konráði að öðru leyti en því að Molbech birtir engin aðkomuorð. Konráð aftur á móti stjörnumerkir allmörg flettiorð á síðunni, en um stjörnuna er þessi skýring í formálanum (1851:iv):

Stjarna (*) er sett a) við þau orð, sem ekki eru gotnesk³ eða ekki eru komin í dönsku úr hinum gotnesku málum. – Þó er henni sleppt við þau, sem orðin eru svo eiginleg dönskunni, að ekki ber á, að þau

3 Skýring í orðabók Konráðs við *gothisk* er 'gotneskur' og henni fylgja dæmin „Den gothiske Folkestamme“ og „En g. Bygning“ (1851:173). Í skýringunni við notkun stjörnunnar (*) er þó ljóst að þar notar Konráð *gotneskur* í merkingunni 'germanskur'.

sjeu útlenzk, t.a.m. *Alter, Præst, Fest, Feber*, – b) við gotnesk orð, sem eru mjög ódönskuleg, t.a.m. *Everlasting, Verruf*.

Í töflu 1 er gerður samanburður á Konráði og fyrirmyndum hans. Í fyrsta dálki eru nefndar flettturnar eins og þær eru í orðabók Molbechs, í öðrum dálki er sýnt hvernig Konráð þýðir eða breytir og í þriðja dálki er sama stafbil úr orðabók Vísindafélagsins.

Christian Molbech	Konráð Gíslason	Videnskabernes Selskab
<i>Deeltagelse</i> , en. <i>det, at tage Deel i noget (saavel egentlig som figurl.) Mange ønskede hans Deeltagelse i Selskabet. Hendes Deeltagelse i min Sorg var oprigtig. Han hørte denne Tidende med stor Deeltagelse.</i>	<i>Deeltagelse</i> , sk., það að <i>deeltage</i> eða <i>tage Deel</i> : <i>Mange ønskede hans D. i dette Forehavende</i> að hann hyrfi að þessu ráði (með þeim). <i>Hendes D. i min Sorg var oprigtig</i> hún samhryggðist mjer einlæglega. <i>Han hørte denne Tidende med stor D.</i> Þessi tíðindi fengu mikið á hann.	<i>Deeltagelse</i> (en) n. s. <i>Gierningen at tage Deel i noget. Han saae det an uden Deeltagellse.</i>
<i>Deeltager</i> , en. pl. -e. <i>En Person, som tager Deel, har Deel i noget.</i>	<i>Deeltager</i> , ft. -e, kk., hlut-takari (Skuggsjá), hlut-takandi (623).	<i>Deeltager</i> (en) n. s. <i>En Person, som tager Deel i noget.</i>
<i>deelviiis</i> , adv. for en <i>Deel</i> , i <i>Henseende til nogle Dele.</i>	<i>deelviiis</i> , ao., í pörtum, (sb. í smásmökkum); að nokkru leyti.	Ekki fletta.
<i>Deels</i> , adv. som altid bruges i en disjunctiv Sætning: for en <i>Deel. Deels</i> var jeg ikke tilstede, og <i>deels</i> var Sagen mig uvedkommende. Besætningen bestod <i>deels</i> af <i>Indfødte, deels</i> af <i>Fremmede.</i> (Jvf. <i>tildeels</i> .)	<i>deels</i> [bls. 79], st., að nokkru leyti, sumpart; <i>Besætningen bestod d. af Indfødte, d. af Fremmede. D. var jeg ikke tilstede, og d. var Sagen mig uvedkommende</i> það var hvorttveggja (eða bæði) ... enda (heldur fornlengt), hvorki ... enda.	<i>Deels</i> , adv. [af <i>Deel</i>]. For en <i>Deel. Bruges meest igientaget, undtagen med præp. til foran.</i>
<i>Deger</i> , et. pl. -e. <i>Et Antal af 10; bruges allene hvor Talen i Handelen er om Huder og Skind.</i>	<i>Deger</i> , ft. -e, kl., [dekur-r], dekur (tugur ⁴ skinna eða húða). b) fjórir dekrar af gerzkri skinnavöru, (sb. timbr Fms. XI, 325).	<i>Deger</i> (et) n. s. [af Lat. <i>Decas</i> .] <i>Et Antal af 10. Bruges allene om Huder og Skind.</i>

4 Konráð notar gleitt letur til áherslu, sbr. formála bls. IV.

Christian Molbech	Konráð Gíslason	Videnskabernes Selskab
<i>Degn, en. pl. -e. [af Lat. Diaconus.] en Kirkebetient, som forestaaer Sangen og opvarter Præsten under Gudstienesten. Degne i Kiøbstæderne have nu i Almindelighed Navn af Klokkere.</i>	<i>Degn, ft. -e, kk., (sb. djákn, já f. ja = e], djákni, meðhjálpari: Degne i Kjøbstæder have nu i Almindelighed Navn af Klokkere.</i>	<i>Degn (en) n. s. Degne [af Lat. Diaconus.] En Mand, som forestaaer Sangen, og opvarter Præsten, i Kirker, hvor der er ingen Klokker. [af Lat. Decanus.] Fordum en Medbroder af et Domkapitel, som kaldtes Domdegn.</i>
<i>Degne, v. n. 1. poet. arte sig som en Degn. „det tidligt degne skal, der gode Degn skal blive.“ Baggesen.</i>	Ekki fletta.	Ekki fletta.

Tafla 1: Samanburður á orðabók Molbechs, Konráðs og orðabók Videnskabernes Selskab.

Ef litið er á einstakar flettur hjá Molbech í töflu 1 má sjá að Konráð fylgir honum nokkuð nákvæmlega. Í þýðingunni á *Deeltagelse* valdi Konráð þó að breyta notkunardæminu, þ.e. „Mange ønskede hans D. i dette Forehavende að hann hyrfi að þessu ráði (með þeim)“ í stað „Mange ønskede hans Deeltagelse i Selskabet“ hjá Molbech. Konráð þýðir notkunardæmin á íslensku notendum til glöggvunar. Í næstu flettu, *Deeltager*, þýðir Konráð ekki skýringu Molbechs en notar orð sem hann þekkir úr fornu máli en örugglega einnig úr samtímamáli. Dæmi eru um bæði orðin *hluttakari* og *hluttakandi* í Ritmálssafni Orðabókar Háskólans allt fram á 20. öld.

Milli orðanna *dearviis* og *Deels* eru 15 orð stjörnumerkt hjá Konráði. Þar er því um tökuorð að ræða og ekki að vænta að þau séu í bók Molbechs (sjá umfjöllun um aðkomuorð í 3.2). Skýringu Molbechs við *Deels* „som altid bruges i en disjunktiv Sætning“ sleppir Konráð en setur þess í stað skammstöfunina st. fyrir „samtenging“. Konráð þýðir ekki notkunardæmin eins og hann gerði undir *Deeltagelse* heldur gefur hann aðeins þýðingu á samtengingunni. Hann breytti röð notkunardæmanna þannig að skýringin `að nokkru leyti, sumpart' á betur við fyrra dæmið en `hvortveggja (eða bæði) ...enda (heldur fornlegt), hvorki ... enda' við hið síðara.

Konráð gerir nokkrar breytingar á flettunni *Deger*. Hann þýðir ekki beint þótt fyrri hluti skýringarinnar sé merkingarlega hinn sami og hjá Molbech og í orðabók Vísindafélagsins og bætir síðan við b) lið `fjórir dekrar af gerzkri skinnavöru'. Þar vísar hann í Fornmannasögur og í orðið *timbr* en það er notað um 40 skinn (Fritzner III:703).

Konráð fylgir Molbech ekki heldur nákvæmlega þegar kemur að orðinu *Degn*. Hann sleppir tilvísun Molbechs til latínu en tengir hins vegar saman orðin *degn* og *djákni* málsögulega. Hann tekur þó með danska notkunardæmið án þess að þýða það en breytir því lítillega, fellir burt greininn í „Kiøbstæderne“. Orðabók Vísindafélagsins skiptir þessari flettu í tvennt, þ.e. eftir því hvort að baki liggur latneska orðið „diaconus“ eða „decanus“. Eins og sjá má fór Konráð ekki þessa leið.

Konráð sleppir að taka með sögnina að *degne* líklegast vegna þess að hún var ekki notuð í almennu máli á hans dögum í Danmörku heldur aðeins í skáldskap. Sú sögn er heldur ekki í orðabók Vísindafélagsins.

Þar sem tengsl við orðabók Vísindafélagsins eru lítil í töflu 1 er látið nægja að bera saman orðabók Molbechs og orðabók Konráðs í töflu 2.

Christian Molbech	Konráð Gíslason
<i>Degnebolig</i> , en. <i>fri Bolig som er bestemt f. en Degn; en Degns Embedsbolig.</i>	<i>Degnebolig</i> djáknahús sem hann hefur leigulaust.
<i>Degneenke</i> , en. <i>Enke efter en Degn.</i>	<i>Degneenke</i> djáknaekkja.
<i>Degnejord</i> , en. <i>J. som ligger til et Degnekald, som hører til en Degns Indkomster.</i>	<i>Degnejord</i> djáknaland sem hann hefur leigulaust.
<i>Degnekald</i> , et. <i>en Degns Embede og Bestilling.</i>	<i>Degnekald</i> djákna-embætti.
<i>Degnekone</i> , en. <i>En Degns Hustru.</i>	<i>Degnekone</i> djáknakona.
<i>Degnekorn</i> , et. <i>En vis aarlig Afgift i Korn af Sognemændene til Degnen.</i>	<i>Degnekorn</i> það korn sem sóknarbændurnir gjalda djáknanum.
<i>Degneoffer</i> , et. <i>Offer i Penge, som gives Degnen ved Høitider.</i>	<i>Degneoffer</i> offer (sb. <i>Offer</i>) sem djákni fær.
<i>Degnepenge</i> , pl. <i>Penge, som gives Degnen ell. Klokkeren i St. for Offer eller Degnekorn. (Holberg.)</i>	<i>Degnepenge</i> , ft., peningar sem djáknanum eru goldnir af sóknarbændum í stað offurs eða korns.
<i>Degnestol</i> , en. <i>Stol i Kirken, hvor Degnen staaer medens han synger.</i>	<i>Degnestol</i> djáknastúka.
<i>Degnetrave</i> , en. <i>En Afgift af visse Neg, som p. nogle Steder gives Degnen i Stedet for Korn i Skieppen, og i Forhold t. Præstens Tiende. (Mandix d. Kammervæsen. 168.)</i>	<i>Degnetrave</i> djáknabundin (ft.), þ. e. kornvöndlar sem djáknum eru goldnir á sumum stöðum í Danmörku í stað korns.

Christian Molbech	Konráð Gíslason
<i>Dei</i> , en. ud. pl. [Isl. <i>Deig</i> ; skrives ogsaa endnu <i>Deig</i> .] en sammenæltet Blanding af Meel og Vand, ell. af Meel, Smør, Melk, m.m. hvoraf bages Brød ell. Kage. (s. <i>Kagedei</i> , <i>Smørdei</i> , <i>Suurdei</i> .) At lægge <i>Dei</i> : blande Meel med Vand og <i>Suurdei</i> eller <i>Giær</i> , og hensætte det for at giæres. At ælte <i>Dei</i> . = <i>Deirand</i> , en 1. Rand af <i>Dei</i> om <i>Kiødmad</i> , som bages. 2. en deiet eller ikke giennembaegt Rand i Brødet.	<i>Dei</i> (líka <i>Deig</i>), án ft., sk., [deig], deig: Lægge <i>D.</i> búa til deig. Ælte <i>D.</i> hnoða deig. Sb. <i>Bolledei</i> , <i>Kagedei</i> , <i>Suurdei</i> .
<i>Deirulle</i> , en. En glathøvet Valse af Træ, som løber om paa en Stok, med to Haandgreb, og hvormed man iævner Deien til visse Slags Bakkelse (<i>Smørdei</i>).	Dálkur 2 á bls. 80: <i>Deirulle</i> eins konar brauðkefli.
<i>Deiskrabe</i> , en. Et Redskab til at skrabe Deien af Truget.	Dálkur 2 á bls. 80: <i>Deiskraber</i> deigskafa.
<i>Deitrug</i> , et. <i>Trug</i> , hvori man ælter <i>Dei</i> (<i>Olufsen</i>).	Dálkur 2 á bls. 80: <i>Deitrug</i> deigtrog (sem deig er hnoðað í).
<i>deiet</i> , adj. 1. blød, som <i>Dei</i> , klæget; ikke giennembaegt. Brødet er <i>deiet</i> . (<i>Moth</i>) 2. besmuurt med <i>Dei</i> .	<i>deiet</i> , ft. <i>deiede</i> , e., sem er eins og deig, hrár (um brauð), illa bakaður: Brødet er <i>d.</i> Sb. <i>klæg</i> . – b) deigugur: <i>Deiede Hænder</i> .

Tafla 2: Frekari samanburður á Molbech og Konráði

Molbech valdi þann kost að setja samsetningar undir grunnorðið eins og oft er gert til að spara rými. Konráð valdi aftur á móti að gera hverja samsetningu að sjálfstæðri flettu. Á eftir orðinu *Degn* (sjá töflu 1) fylgja allmargar samsetningar með *Degne*- að fyrri lið. Hér var farin sú leið að brjóta upp flettugrein Molbechs til þess að einfalda samanburðinn. Konráð þýðir ekki alltaf nákvæmlega eftir Molbech en lagar skýringarnar þannig að þær samræmist betur íslenskum veruleika. T.d. tekur hann fram við *Degnetrave* að það eigi við sums staðar í Danmörku.

Konráð stytir talsvert texta Molbechs við orðið *Dei*. Hann hefur líklega talið að allir vissu hvað þyrfi til að búa til deig. Öll þrjú orðin sem vísað er til, *Bolledei*, *Kagedei*, *Suurdei* hjá Molbech og Konráði, eru skýrð sérstaklega á réttum stað í stafrófsröðinni. Samsetningarnar *deirulle*, *deiskrabe* og *deitrug* eru sett eftir reglu Molbechs undir orðið *dei*. Samsvarandi orð hjá Konráði lenda samkvæmt stafrófsröðinni í öðrum dálki á síðu 80 en eru þó tekin með hér til samanburðar við Molbech. Flettunni *deiet* breytir Konráð lítillaga og bætir við notkunardæmi í b) lið um „deigugar hendur“ sem einnig er að finna í orðabók Vísindafélagsins og er hugsanlega fengið þaðan.

Af þessum samanburði á einum dálki í orðabók Konráðs í töflum 1 og 2 má sjá hvernig hann fylgir Molbech. Sömu flettiorð eru valin og flest notkunardæmin einnig. Konráð breytir helst þar sem hann telur að önnur skýring eða viðbótarskýring komi hans notendum að meira gagni. Hann styttir aðeins þar sem hann telur lýsingu Molbechs of langa (*Degn, Dei*) og óþarfa fyrir tvímála bók. Eins og sjá má af þriðja dálki í töflu 1 er meiri munur á skýringum í orðabók Vísindafélagsins. Konráð hefur væntanlega haft útkomin bindi við höndina við verk sitt þótt hann hafi kosið bók Molbechs sem meginfyrirmynd. Hún hentaði einnig betur en þau sex bindi sem komin voru út af orðabók Vísindafélagsins.

3.2 Samanburður við orðabók Meyers

Eins og Konráð getur um í formála notar hann orðabók Ludwigs Meyers frá 1844 við val á þeim orðum sem ekki eru „gotnesk“ (sbr. 3.1). Öll stjörnumerktu orðin í fyrsta dálki á síðu 80 hjá Konráði er að finna í bók Meyers. Konráð notar orðið „aðkomuorð“ í formálanum að dönsku orðabókinni (1851:IV) þar sem hann segir: „Ekki hef jeg síður vísað til íslenzkra aðkomuorða, þegar dönsk orð eru dregin af þeim, (sb. Eventyr, Præst).“ Hann virðist, að mínu mati, greina á milli fornra tökuorða, sem sameiginleg eru norrænum málum, og þeirra sem ný eða nýleg eru í dönsku.

Konráð Gíslason	Ludvig Meyer
* <i>defect</i> , ft. -e, e., sem í vantar (t.a.m. um bók)	Defect , l., en, Mangel, Feil, Ufuldstændighed, det Manglende, f. Ex. i en Kasse, i en Bog ...
* <i>Defension</i> , ft. -er, sk. vörn, einkanl. máls-vörn	Defension , l., en, Forsvar, Modværge, Værn, Forsvarslinie, Forsvarsanstalt, Forsvarsskrift
* <i>defensiv</i> , ft. -e, e., varnar-. Gagnstætt: <i>offensiv</i> .	defensional, defensionel og defensiv , beskyttende, værnende, forsvarrende
* <i>Defensivalliance</i> = <i>Forsvarsforbund</i> .	Defensiv-Alliance , Forsvarsforbund
* <i>Defensivkrig</i> = <i>Forsvarskrig</i> .	D=Krig, Forsvarskrig
* <i>Defensivvaaben</i> = <i>Forsvarsvaaben</i> .	D=Vaaben, Forsvarsvaaben
* <i>Defensor</i> , ft. -er, kk., sakar verjandi.	Defensor , l., en, Forsvarer, Beskytter, Forfægter; den Sagfører, som i offentlige Sager gives den Tiltalte til Forsvar

Konráð Gíslason	Ludvig Meyer
* <i>Deficit</i> (eiginl. sögn, og merkir vantar), kl., það sem á vantar að tekjurnar jafnist við gjöldin.	Deficit , l., et, = Defect , f.d., is. det, som mangler i Indtægten, for at gjøre den lige med Udgiften
* <i>Defilee</i> , sk. þröngur vegur, Sb. <i>Huulvei</i> .	Défilé , fr., en, snever Pas, Huulvei
* <i>defilere</i> , h. s. 1, (um herlið) fara þröngvan veg; ganga fram hjá e-u hver eptir öðrum.	défilere , drage gennem en Huulvei; trække forbi i smalle Rækker el. ogsaa Mand for Mand
* <i>definere</i> , ás. 1, nema (þ.e. kveða svo skýrt á hvað e-ð sje að maður greini það nægilega frá öllu öðru.)	definere , l., eg. angive Grændserne; bestemme, give et nøie bestemt Begreb om en Ting
* <i>Definition</i> , ft. -er, sk. nám (sb. <i>definere</i>).	Definition , en, Udvikling, Forklaring, Bestemmelse (af et Begreb)
* <i>definitiv</i> , ft. -e, e., = <i>endelig</i> b. Sb. <i>uforanderlig</i> , <i>ugjenkaldelig</i> .	definitiv ; afgjørende; endelig afgjort, bestemt (modsat provisorisk), uforanderlig, uigjenkaldelig
* <i>Degeneration</i> , sk., það að <i>degenerere</i> .	Degeneration , en, Udartning, Vanslægtning
* <i>degenerere</i> , h. s. 1, úrkynjast, affeðrast, verða ættleri.	degenerere , udarte, vanslægte, blive slettere
* <i>Degradation</i> , ft. -er, sk., það að <i>degradere</i> .	Degradation , ny l., en, Afsættelse fra en Værdighed, Fornedrelse, Nedværdigelse
* <i>degradere</i> , ás. 1, setja e-n í lægri stjett úr annari hærri (sb. lækka).	degradere , l., nedsætte, afsætte fra en Værdighed, fornedre, nedværdige

Tafla 3: Samanburður á orðabók Konráðs og orðabók Meyers

Samheiti eru talsvert notuð hjá Meyer og virðist Konráð kjósa þá leið að velja algengustu merkinguna en láta vera að finna viðeigandi þýðingar á öllum þeim orðum sem Meyer hafði valið sem skýringu. Hann nýtir sér því bók Meyers á annan hátt en Molbechs eins og sjá má í töflum 1 og 2. Sem dæmi má nefna fyrsta orðið *defect*. Hjá Meyer er skýringin 'Mangel, Feil, Ufuldstændighed, det Manglende, f. Ex. i en Kasse, i en Bog.' Konráð notar aðeins síðasta hluta skýringarinnar. Sama gildir t.d. um **Defension* sem Konráð þýðir með „vörn, einkanl. Málsvörn“ en Meyer með „Forsvar, Modværge, Værn, Forsvarslinie, Forsvarsanstalt, Forsvarsskrift“. Við sagnleidd nafnorð, t.d. *Definition*, *Degradation*, lætur Konráð sér nægja að nota „það er ...“ og vísa til sagnarinnar.

Í formálanum að *Danskri orðabók með íslenskum þýðingum* minnst Konráð ekkert á hvernig hann kaus að nýta sér orðabók Meyers (sbr. 3.1 hér að ofan). Stjörnumerkingin (*) framan við orð á þó við þau sem hann sótti til Meyers án þess að það sé nefnt sérstaklega.

3.3 Aðrar heimildir Konráðs

Þegar Konráð vann við orðabókina hafði hann einkum tvær íslenskar orðabækur til þess að fletta í. Önnur var íslensk-latnesk-dönsk orðabók Björns Halldórssonar frá 1814 og hin dönsk-íslensk orðabók Gunnlaugs Oddssonar frá 1819. Fremst í orðabók Konráðs er listi yfir skammstafanir og tilvísanir og þar eru þessar orðabækur báðar nefndar sem heimildir en að auki fjölmörg fornrit, *Um frumparta íslenzkrar tungu í fornöld* eftir Konráð sjálfan (1846), *Grasafræði* Odds Hjaltalíns (1830), *Jóns Péturssonar Lækninga-Bók* (1834), orðabókarhandrit Páls Þorbergssonar og *Safn af íslenzkum orðskviðum* sem Guðmundur Jónsson gaf út (1830). Að auki tilgreinir hann sem heimildarmenn Björn Gunnlaugsson, Bjarna Thorarensen, Hallgrím Scheving, Jónas Hallgrímsson og Jóhann Halldórsson sem áður var getið.

Ekki er að sjá að Konráð hafi haldið fram orðasmíð Jónasar Hallgrímssonar, vinar síns, því að nokkur leit er að skammstöfuninni J. H. í bókinni. Hana fann ég þó á nokkrum stöðum, m.a. við þýðinguna á danska orðinu *giftblanderske* sem er 'eiturbyrila'. Það notar Jónas í þýðingu á kvæðinu *Dagrúnarharmur* (Die Kindsmörderin) eftir Friedrich Schiller: „sætt er eitur þitt, / er það nú goldið, / eiturbyrila / ungra hjartna“ (Jónas Hallgrímsson 1989:80). *Giftblander* er fletta í orðabók Molbechs en Konráð gerir úr tvær:

Molbech:

Giftblander, en. den, som blander, tilbereder G. i den Hensigt dermed at dræbe Andre. Om et Fruentimmer bruges: Giftblanderske, en.

Konráð:

Giftblander, ft. -e, kk., sá sem byrlar eitur.

Giftblanderske, ft. -r, kvk., sú sem byrlar eitur, (eiturbyrila, J. H.).

Jónasar er einnig getið við orðið *Æther*. Þar velur Konráð skýringarnar 'eldloft', sem hann hefur úr orðabók Björns Halldórssonar, 'vindblá' úr orðabók Gunnlaugs Oddssonar og 'ljós-vaki' sem Jónas hafði notað í þýðingu sinni á *Stjörnufræði* G. F. Ursins (1842).

Við orðin *Skildpadde*, *Bløddyr* og *Hvirveldyr* notar Konráð 'skjaldbaka', 'lindyr', 'hryggdýr' án þess að geta Jónasar sem fyrstur notaði þau á prenti í grein í *Fjölni* hálfum öðrum áratug áður (1836:2,3) en hugsanlegt er að orðin hafi þegar verið það algeng í málinu að honum hafi þótt óþarft að geta heimildarinnar. Konráð er einnig með orðið 'skelpadda' sem þýðingu á *skildpadde* en það er ekki komið frá Jónasi

heldur er talsvert eldra og a.m.k. frá miðri 17. öld (sbr. Ritmálssafn Orðabókar Háskólans). Orðin *Bløddyr* og *Skildpadde* eru bæði flettiorð hjá Molbech sem lýsir þeim allnákvæmlega. *Hvirveldyr* er ekki fletta hjá honum.

Oft hefur verið vitnað til orðasmíðar Jónasar í þýðingu hans á *Stjörnufræði* Urins. Ég skoðaði þau orð sem Bjarni Vilhjálmsson tók saman í grein í *Skírni* um nýyrði Jónasar í *Stjörnufræðinni* (1944: 99–130) og bar saman við orðabók Konráðs til þess að kanna hvort Konráð hefði leitað í smiðju til Jónasar. Ég skoðaði einnig orðabók Gunnlaugs Oddssonar frá 1819 sem reyndar hafði ekki öll orðin sem flettu. Eins og við mátti búast eru ekki nærri öll orðin í bók Molbechs þar sem hann sneiddi hjá aðkomuorðum. Tafla 4 sýnir samanburðinn. Í fremsta dálki eru dönsk orð úr orðabók Konráðs, í næsta dálki skýringar hans, í þriðja dálki orð frá Jónasi samkvæmt grein Bjarna Vilhjálmssonar og í síðasta dálki eru skýringar orðanna í orðabók Gunnlaugs. (+) er settur við þau orð sem eru í orðabók Molbechs.

Danska	Konráð Gíslason	Jónas Hallgrímsson	Gunnlaugur Oddsson
Apogæum	fjarlægð við jörðina	jarðfjærð	
Atmosphære	gufuhvolff, þokulopt	lofthaf	dampahvolff
Brændeglas (+)	eldgler	eldgler	
Centrifugalkraft	frásóknarafl hlutanna ⁵ frá miðpunkti hringbrautar sinnar	miðflóttaafl	midpunktflýandi-kraftur
Centripetalkraft	tilsóknarafl hlutanna til miðpunkts hringbrautar sinnar	miðsóknarafl	midpunkt-sækiandi kraftur
Elastisk (+)	sem lætur undan og tekur við sjer aptur, sem teygja eða staður er í	fjaðurmagnaður	hefir spennikraft
Electricitet (+)	rafmagn	rafurmagn	
Ellipse	2) (í mælingafræði), sporbaugur	sporbaugur	
Friction	núningur, nugg	núningsfyrirstaða	nagg, nudd, vidnúningur
Mikroskop	stækkunargler, stækkunarpípa	sjónauki	stækkunargler
Nutation	rugg, eða þvíuml.	rugg	

5 Engar skýringar eru á því í formála hvers vegna Konráð kys að nota tvenns konar leturstærð í skýringum sínum (sjá einnig töflu 1 og 2).

Danska	Konráð Gíslason	Jónas Hallgrímsson	Gunnlaugur Oddsson
Optik (+)	sjónarfræði, ljósfræði	ljósfræði	siónar-liós-liós-striála-fræði
Passat (+)	stöðuvindur (í nánd við miðjarðarhring)	staðvindur	þrávindi, rakinn vindr frá sømu átt (í sudur hluta heims)
Pendul	dingull (t. a. m. á stunda-klukku; sb. <i>Uro b</i>)	hengill	veifilód
Perigæum	mest nálægð (reikandi stjörnu) við jörðina	jarðnánd	
Perturbation	truflun	hrakningar	
Sextant	mælingaverkfæri sem skipt er í (sjöttung hrings =) 60 mælistiga	sextungur, skuggsjárbaugur	
Synsvinkel (+)	sjónarhorn (þ. e. kverk á milli tveggja hinna yztu sjónargeisla	sjónarhorn	
vandret (+)	marflaturlatur (B. H.)	hafjafn	
Vaterpas (+)	lágmaal (Sch.)	hafflatarmælir	marfløtismælir
Æther	eldloft, vindblá (B. H., G. O.), ljós-vaki (J.H.)	ljósvaki	vindblá, heid

Tafla 4: Samanburður á notkun nýyrða hjá Konráði, Jónasi og Gunnlaugi

Tvö orðanna eru umrituð hjá Konráði: *elastisk* og *sextant*. *Elastisk* skýrir hann: 'sem lætur undan og tekur við sjer aptur, sem teygja eða staður er í.' Molbech skýrir orðið: 'som har den Egenskab, efter en udvidende Sammentrykning ell. Bøining igien at trække sig sammen og indtage sit forrige Rum'. *Sextant* skýrir Konráð: 'mælingaverkfæri sem skipt er í (sjöttung hrings =) 60 mælistiga.' Það orð er ekki hjá Molbech og Konráð hefur skýringuna ekki heldur úr orðabók Vísindafélagsins. Reikna má með að hann hafi leitað víðar fanga en í þeim þremur orðabókum dönskum sem hann tilgreinir enda segir hann í formála að hann hafi tekið auk orðabókanna áður nefndu „fátt eitt úr öðrum bókum eða tali“ (1851:III). Áhugavert er að skoða orðið *electricitet*. Molbech hefur lýsingu á fyrirbærinu: 'Den Tilstand hos visse Naturlegemer, hvorved den electricke Materie bliver virksom, eller ytrer sig i electricke Phænomener'. Konráð sleppir henni en notar aðeins jafnheitið 'rafmagn'. Jónas hafði notað *rafurmagn* í þýðingu sinni. Elsta dæmi Orðabókar Háskólans um *rafmagn* er úr riti Gísla Brynjólfssonar, *Dagbók í Höfn*, sem skrifuð var 1848 (1952:260).

Vel má hugsa sér Konráð hafi orðið annaðhvort frá honum eða öðrum Íslendingum samtímis í Kaupmannahöfn.

Eins og sjá má notaði Konráð fimm af nýyrðum Jónasar, *sporbaugur*, *eldgler*, *ljósfræði*, *ljósvaki* og *sjónarhorn*, en getur hans aðeins við orðið *ljósvaki* þar sem hann er einnig með tillögur Björns Halldórssonar og Gunnlaugs Oddssonar. Frá Gunnlaugi hefur hann orðið *vindblá* en jafnheitið 'stækkunargler' sem Gunnlaugur gefur einnig á *Mikroskop* er ekki frá honum komið því að það þekkist að minnsta kosti frá síðasta þriðjungi 18. aldar (sbr. Ritmálsskrá Orðabókar Háskólans).

3.4 Viðtökur og ummæli

Ekki virðist mikið hafa verið fjallað um *Danska orðabók með íslenskum þýðingum* á Íslandi þegar hún kom út. Við allnokkra leit í íslenskum samtíma ritum og blöðum hef ég ekki getað fundið eiginlegan ritdóm. Árið 1855 var dreift boðsbréfi sem eignað er Jóni Ingjaldssyni presti á Húsavík þar sem hann gagnrýnir bókina og boðar viðauka með nánari skýringum. Þeir komu þó aldrei út.

Eins og fram hefur komið ritaði Björn M. Ólsen eftirmæli eftir Konráð látinn og kom þar m.a. að dönsku orðabókinni (1891: 65–66). Þar segir hann að það hafi einkum verið áhugi Konráðs að hreinsa dönskuslettur úr málinu sem hvatt hafi hann til þess að vinna orðabókina. Honum hafi gramist að lærðir menn hugsuðu á dönsku „og klæddu því hugsanir sínar í danskan búning, jafnvel þegar þeir töluðu eða skrifuðu á íslensku“ (66). Ekkert kemur fram í formála Konráðs né í öðrum skrifum, sem varðveitt eru og ég hef farið yfir, hvers vegna þeir Jón Halldórsson ákváðu að fara af stað með verk af þessu tagi. Björn M. Ólsen var góður vinur Konráðs, samtíma honum í Kaupmannahöfn um skeið á þeim tíma sem hann vann að dönsku orðabókinni og íslensk-ensku orðabókinni (sbr. 2. kafla), og er ekki ólíklegt að hana hafi borið á góma í samtölum þeirra. Ekkert er þó til skjalfest um það svo vitað sé.

Á sömu síðu hélt Björn M. Ólsen áfram:

Engin bók, sem út hefur komið á þessari öld, hefur átt jafnmikinn þátt í að útríma dönskublendingnum úr tungu vorri eins og þessi. Áhrif hennar á ritmál vort hafa verið fjarska mikil og væri hægt að sanna það með mörgum dæmum, ef tími og rúm leifði. Hún hefur nú í 40 ár kent ungum og gömlum eigi að eins að skilja dönsku heldur og að þíða hana á hreina íslensku. Þíðingar orðanna eru vanalega smellnar og rammíslenskar, og þó að það komi stundum fyrir, að það

finnist ekki í orðabókinni, sem verið er að leita að í þann svipinn, þá má segja slíkt hið sama um allar orðabækur.

Af þessari umsögn má sjá að Björn M. Ólsen kunni að meta orðabók Konráðs. Ekki er vitað hve stórt upplagið var en bókin mun hafa verið til víða á heimilum, einkum þeim efnameiri. Umdeilanlegt er hvað er smellið og hvað ekki, en afar góð þekking Konráðs á fornu máli hefur komið að góðu haldi við skýringarnar. Þótt Konráð hafi búið erlendis mestan hluta ævinnar var hann í góðu sambandi við Íslendinga í Kaupmannahöfn og var þannig stöðugt í tengslum við samtímamálið og gat leitað til heimildarmanna ef þörf var á.

Konráð leit sjálfur á orðabók sína sem frumsmíð „en frumsmíðir standa til umbóta“ (1851: IV). Það kom í hlut Jónasar Jónssonar á Hrafnagili að semja og gefa út ásamt fleirum næstu dansk-íslenska orðabók, undir titlinum *Ný dönsk orðabók með íslenskum þýðingum*. Hún var gefin út í Reykjavík 1896. Frumrit Jónasar er varðveitt í eiginhandarriti á Amtsbókasafninu á Akureyri. Í formála fyrir eiginhandarritinu frá ágúst 1896 kemur fram að bókin var frá Jónasar hendi að mestu unnin á árunum 1885–1889. Þá var Konráð enn á lífi (d. 1891) en ekki er vitað til þess að Jónas hafi verið í sambandi við Konráð. Meðal rita sem Jónas studdist við var „Konráðs orðabók“. Orðabók Jónasar kom þó ekki út fyrr en 1896 og höfðu þá verið gerðar ýmsar viðbætur og leiðréttingar.

Formáli Jónasar, sem hann skrifaði 1896, var ekki notaður í útgefnu bókinni heldur formáli, sem undirritaður er af Birni Jónssyni, einum meðhöfundi bókarinnar á síðari stigum, einnig í ágúst 1896. Þar kemur fram að Íslendinga hafi lengi vantað „nýja, handhæga, alþýðlega danska orðabók með íslenskum þýðingum“. Íslenska og danska hafi breyst mikið frá því að orðabók Konráðs kom út, orð dáíð, önnur orð breytt um merkingu og ný orðið til. Í nýju bókinni sé mikið efni sem ekki sé í bók Konráðs „en vantar yfirleitt það eitt af hennar orðaforða, sem nú er úrelt orðið eða þar um bil“ (1896: III). Um bók Konráðs segir enn fremur: „Það er einnig óhandhæg bók og ekki við alþýðu hæfi, meðal annars nokkuð kostnaðarsöm.“ Vissulega var nýja bókin handhægari, ódýrari og nær nútíma og ekki hefur þótt eftir útkomu hennar ástæða til að gefa bók Konráðs út að nýju.

Orð Björns Jónssonar lýsa að einhverju leyti viðhorfi samtíma hans til *Danskrar orðabókar með íslenskum þýðingum* þótt hann sé um leið að mæla með nýju bókinni. Hún þótti að sögn Björns dýr og hefur það því helst verið á færi hinna efnameiri eða þeirra sem nauðsynlega þurftu á orðabók að halda að eignast hana.

4 Önnur orðabókaverk

Í áður nefndu bréfi til konungs frá 1847 (sjá 1. kafla) kom fram að Konráð ynni að orðabók yfir íslenskt nútímamál með dönskum skýringum. Sennilega er það þetta verk sem Hallgrímur Scheving talar um í bréfi til Konráðs 1849 en þar stendur m.a. (Finnbogi Guðmundsson 1970:190–191):

Þér mælist til, að eg feli yður á hendur mín íslenzku orðasöfn, af því „Samfundet for dansk Litteraturs Fremme“ hafi beðið yður að taka saman íslenzka og danska orðabók yfir nýrra málið. Segizt þér síðan láta bókina koma út undir mínu nafni, en setja merki við það sem þér bætið við. ... skal eg með sumarskipum senda yður smásaman nokkra stafi af söfnum mínum, eftir því sem þér í hvert sinn látið mig vita hvað því verki líður.

Samfundet for dansk Litteraturs Fremme hefur leitað til Konráðs þegar ljóst var að dansk-íslenskri orðabók yrði senn lokið. Konráð var í góðu vinfengi við Hallgrím Scheving og hafði m.a. unnið við að hreinrita orðasafn hans (Guðrún Kvaran 2008:156). Engar heimildir eru til um að Hallgrímur hafi sent Konráði efni til þessarar orðabókar.

Svipuð voru örlög ensk-íslenskrar orðabókar. Af gögnum Konráðs að ráða, sem varðveitt eru á Stofnun Árna Magnússonar í Kaupmannahöfn, virðist hann hafa ætlað sér að semja slíka orðabók en lítið orðið úr framkvæmdum. Til eru orðalistar og drög að enskri málfræði en engin drög að orðabók. Hafi hann farið af stað með verkið virðist ekkert hafa varðveist af því.

Enn eitt orðabókarverk virðist með öllu týnt sem þó er vitað að Konráð vann að um árabíl. Þar er um að ræða orðabók yfir forna málið með dönskum skýringum. Hann fékk til hennar styrk frá danska Kirkju- og kennslumálaráðuneytinu á árunum 1857–1863 og sótti um viðbótarstyrk 1865 til að búa bókina til prentunar. Rakið verður hér það sem vitað er um þetta verk til að varpa ljósi á þessa miklu fyrirætlun Konráðs.

Ýmis bréf eru til sem styðja það að Konráð vann við fornnorræna orðabók með dönskum skýringum og var kominn nokkuð áleiðis. Hann sótti um styrk til „Det høie Ministerium for Kirke- og Underviisningsvæsenet“ í bréfi frá 22. maí árið 1856⁶ til að vinna „en muligst fuldstændig oldnordisk-dansk Ordbog“. Hann telur sjálfan sig vel til

6 Öll bréfin sem vitnað er til í þessum kafla eru varðveitt í Konunglega bókasafninu í Kaupmannahöfn en ljósrit af þeim eru í mínum fórum.

Þess fallinn að vinna slíkt verk þar sem hann búi að langri reynslu af lestri fornra texta og reiknar með að þurfa til verksins sex ár. Í bréfinu er ítarleg lýsing á því hvað vinna þurfi. Í fyrsta lagi þurfi að orðtaka ýmist að nýju eða að hluta að nýju það elsta og gagnlegasta af handritum og skjölum í Kaupmannahöfn, Stokkhólmi og Uppsölum en einnig í Kristjaníu og Wolfenbüttel. Í öðru lagi þurfi að fara yfir þau orðfræðilegu verk yfir forna málið sem unnið er að, bæta í þau og endurskrifa. Þar nefnir hann einungis Sveinbjörn Egilsson og safn í eigu sjálfs sín. Í þriðja lagi þurfi að fella saman þau tvö söfn sem nefnd séu undir lið tvö. Með því að hefja verkið á 400 stærstu orðunum, lýsa merkingarþróun þeirra á kerfisbundinn hátt og sýna mikinn fjölda dæma með skýringum og þýðingum verði hægt að létta þá vinnu sem eftir er og spara með því tíma og fé. Konráð taldi nauðsynlegt að ráða aðstoðarmenn til orðtöku og uppskrifta og eins að fylgjast vel með því sem verið væri að vinna í orðabókagerð í öðrum löndum. Sótti hann um 1500 ríkisdali.

Eins og fram hefur komið fékk Konráð jákvætt svar við styrkumsókninni með bréfi frá 11. júlí 1857 og 3. mars árið 1858 skrifaði hann bréf til þess að gera grein fyrir því sem unnist hefði á fyrsta styrkárinu. Í bréfinu er ítarleg verklýsing og þar kemur fram í lokin að engin grein sé endanlega frágengin. Það efni sem safnað hafi verið til verksins sé mikið og reiknað sé með að um 15.000 greinar verði til í lok næsta fjárlagaárs 1858/59.

Í bréfi dagsettu 2. apríl 1859 sendir hann umbeðið ársyfirlit yfir verkið. 12.000 greinar höfðu verið unnar, mikilvæg handrit orðtekin og mikil vinna lögð í að undirbúa næsta skammt af greinum. Vonast Konráð til að verða hálfnaður með verkið um mitt ár 1860. Í bréfi frá 31. mars 1860 telur hann sig hafa náð settu marki, tilbúnar greinar séu nú um 20.000 og þegar sé vinna hafin við margar aðrar. Í lokin skrifar hann:

For at fatte mig kort, skal jeg hertil kun föie Forsikringen om, at jeg opfylder en i hver Henseende behagelig Pligt, idet jeg af al min Evne arbejder hen til det Maal snarest muligt at fuldende en oldnordisk-dansk Ordbog der kunde bidrage ikke allene til at udvide den hele Synskreds i dette Fag, men ogsaa til at berigtige vildtfarene og vildtledende Opfatninger.

Samkvæmt bréfi frá 30. mars 1861 telur hann sig tilbúinn með 30.000 greinar, helminginn af fyrirhuguðu verki, og að fjölmargar greinar til viðbótar séu nær fullskrifaðar. Hann sér ekkert því til fyrirstöðu að hægt verði að prenta ritið í lok þess tíma sem fyrirhugaður var til

vinnunnar. Næsta bréf er skrifað 1. apríl 1862. Kaflinn um orðabókina er stuttur, tíu handskrifaðar línur sem enda þannig:

troer jeg nu at kunne indskrænke mig til at udtale Haabet om i det tilstundende Finantsaar at faae dette Værk – enten ganske, eller saa godt som – færdigt i Manuscript.

11. mars 1863 skrifar Konráð enn bréf um stöðu verksins og sækir um áframhaldandi styrk á fjárlagaárinu 1863–1864. Þar bendir hann á að 18 arkir af orðabók Johans Fritznars séu þegar komnar út með greinum frá *a – hrímdrif* og væntanleg sé á næstu vikum orðabók Eiríks Jónssonar. Þessi tvö verk ásamt orðabók Sveinbjarnar Egilssonar muni án efa bæta tímabundið þann skort sem sé á fornnorrænum orðabókum. Sjálfur hafi hann því ákveðið að auka sitt verk með því að bæta við mikilvægum flokki orða sem í séu öll norræn viðurnefni og eiginnöfn og val af landræðilegum nöfnum. Í bréfi frá 16. mars 1864 nefnir hann aftur að hann hafi ákveðið að stækka sína væntanlegu orðabók og telur sig þurfa tvö ár til viðbótar til þess. Prentun ætti að geta hafist 1. apríl 1865. Óskar hann eftir 1000 ríkisdala styrk á næsta fjárlagaári.

Síðasta bréfið til ráðuneytisins í þessari lotu er frá 6. febrúar 1865. Þar gerir Konráð rækilega grein fyrir því hvernig hann hyggist stækka verk sitt með áður nefndum viðbótum og sækir aftur um 1000 ríkisdala styrk til að búa það endanlega til prentunar.

Nú virðist hafa orðið hlé í vinnu við forníslensk-danska orðabók en í bréfi dagsettu 27. september 1872 til kirkju- og menntamálaráðuneytisins gerir Konráð grein fyrir þessari töf:

Det vil være i det høje Ministeriums Erindring, at Regeringen i sin Tid i en Række Aar (1856 fgg.) ydede mig Understøttelse til Udarbejdelsen af en oldnordisk-dansk Ordbog. At jeg ikke kom til Ende hermed, laa til Dels i de litterære Forhold paa den oldnordiske Lexikografis Omraade.

Síðan getur hann þess að frá 1860 hefðu komið út fimm orðabækur yfir íslenskt mál sem hann nefnir allar og fer um gagnrýnisorðum. Þetta voru *Lexicon poëticum* Sveinbjarnar Egilssonar frá 1860 („et særdeles fortjenstligt Arbejde, der bragte Indsigten i oldnordisk Poesi et stort Skridt fremad, er nu efter min Anskuelse for en stor Del forældet.“), *Oldnordisk ordbog* Eiríks Jónssonar frá 1863 („er udarbejdet efter en meget utilfrædsstillende Plan, som blev forfatteren ligesom 'paaoktrojeret'“), *Altnordisches Glossar* eftir Möbius frá 1866 („Men det ville dog være urimeligt her at forlange en fuldkommen Nøjagtighed

i det enkelte.“), fornmálsorðabók Johans Fritzners frá 1867 („har sine ubestridelige Fortjenester, men gjør sig paa den anden Side skyldig i mange Misforstaaelser.“) og orðabók Cleasbys sem farið var að gefa út í heftum. Hann gagnrýndi þær allar, sagði þær þegar úreltar og ekki fræðilega unnar og taldi, þrátt fyrir nýlegar fornmálsorðabækur, fulla þörf á þeirri sem hann hafi unnið að um árabíl. Athugasemdin um orðabók Cleasbys er birt í kafla 2.

Konráð telur allar orðabækurnar fimm aðeins „Glossarier – kun Brudstykker af et fuldstændigt Lexikon.“ Hann segist engan veginn hafa gefið verkið upp á bátinn og muni helga því allan sinn tíma þegar hann hafi komið frá sér útgáfu á Njálu.⁷ Síðan bætir hann við:

Vejen til et Maal, man ønsker at naa, synes gjerne kortere, end den er i Virkeligheden; og da jeg fattede det Forsæt at udarbejde en Ordbog over en saa rig Litteratur, som den oldnordiske, havde jeg – skjönt øvet i lexikalsk Excerpting omtrent fra mit 18. Aar af – dog ikke rigtig udmaalt Forholdet imellem min Opgave og den Tid det vilde være muligt for en enkelt at anvende derpaa i nogle Aar.

Hann óskar eftir því að fá að ráða Eirík Jónsson sem aðstoðarmann þar sem honum sé ljóst að þetta sé ekki eins manns verk. Nú reiknar hann með að 10–12 ár þurfi til að ljúka verkinu sem sé allt byggt á úrvinnslu úr handritum sem séu mjög mörg og sum erfið viðureignar.

Fornmálsorðabók Konráðs kom aldrei út og handritið virðist með öllu glatað. Sama er að segja um svar ráðuneytisins um viðbótarstyrk. Engin bréf frá Konráði til ráðuneytisins eru varðveitt yngri en bréfið frá 1872. Hugsanlegt er að ýmsir orðalistar og listar yfir örnefni, viðurnefni og mannanöfn, sem varðveittir eru í Stofnun Árna Magnússonar í Kaupmannahöfn, hafi orðið til eftir að hann ákvað að hafa með slík nöfn án þess að unnt sé að fullyrða það en búast hefði mátt við að meira efni lægi fyrir miðað við að Konráð hafði unnið við verkið frá 1857 til 1865 þegar hann sækir um útgáfustyrk og virðist enn vera að 1872.

5 Niðurlag

Af því sem fram er komið má ljóst vera að Konráð hafði áhuga á að Íslendingar eignuðust orðabækur. Annars hefði hann tæplega eytt jafn miklum tíma í verk af því tagi. Hann virðist hafa haft ýmislegt á þrjónunum sem ekkert varð úr eins og íslensk-danska nútímamáls-

⁷ Bókin kom út 1875.

orðabók og ensk-íslenska orðabók. Aftur á móti er ástæðulaust að rengja að hann hafi unnið að forníslensk-danskri orðabók þótt handritið virðist hafa glatast. Hann setti markið hátt, sennilega of hátt fyrir einn mann sem var önnur kafinn við kennslu og rannsóknir. Ekkert bendir til að Eiríkur Jónsson hafi komið að verkinu með honum. Skýrslurnar sem varðveist hafa eru áhugaverð heimild um fyrirhugað verk. Metnaður Konráðs var mikill, allt skyldi vera með en markmiðið var engan veginn raunsætt fyrir einn mann,

Vinna við orðabók Cleasbys var einnig staðreynd þótt ekki sé unnt að benda á þátt Konráðs eða annarra Íslendinga í því verki sem út kom 1874. Eins og fram kom í öðrum kafla segist Konráð þó hafa skipulagt verkið, haft umsjón með því og unnið mikið af því efni sem lá fyrir þegar bókin kom út.

Eina bókin sem Konráð fylgdi til loka 1851 var *Dönsk orðabók með íslenskum þýðingum* sem hann vann í sama anda og Molbech að skrifa fremur merkingarlýsingar og sýna notkun með dæmum en að skýra með jafnheitum þótt vissulega séu jafnheiti notuð í bland. Molbech gaf út einmála orðabók. Í slíkri orðabók er aðeins eitt umfjöllunarmál sem yfirleitt er móðurmál notendanna sem bókin er ætluð. Molbech skrifaði fyrir danska notendur eða notendur sem gott vald höfðu á dönsku (*Nordisk leksikografisk ordbog* 1997:111). Konráð skrifaði hins vegar fyrir íslenska notendur og hans bók er tvímála. Í tvímála orðabók eru tvö umfjöllunarmál, annars vegar viðfangsmálið, sem er málið á flettiorðaskránni, og hins vegar markmálið sem yfirleitt er móðurmál höfundar eða mál sem hann ræður mjög vel við (*Nordisk leksikografisk ordbog* 1997:257–258). Skýringar í tvímála orðabókum eru oftast jafnheiti en í töflum 1 og 2 má einnig sjá þess dæmi að Konráð þýðir skýringar Molbechs. Í töflu 3 má sjá að Konráð notar eingöngu jafnheiti þegar hann nýtir sér orðabók Meyers.

Ýmsir hafa nefnt að *Dönsk orðabók með íslenskum þýðingum* hafi haft mikil áhrif til hins betra á málnotkun manna. Áður var vitnað í ummæli Björns M. Ólsen en aðrir sem tjáð hafa sig um áhrifin eru Jónas Kristjánsson (1986:143), Jakob Benediktsson (1969: 105) og Kjartan Ottósson. Kjartan nefndi þó að orðabókin hefði lítið stuðlað að aukinni útbreiðslu nýyrða (1990:101). Enginn þessara þriggja styðst við athugun á áhrifum bókarinnar og er mér ekki kunnugt um að slík athugun hafi farið fram.

Engar heimildir hafa fundist um að Konráð hafi haft endurútgáfu í huga þótt hann hafi lifað í fjörutíu ár eftir að bókin kom út. Öll hans elja á orðabókarsviðinu eftir að *Danskri orðabók með íslenskum*

Þýðingum lauk fór í fornmálsorðabókina með dönskum skýringum. Þegar orðabók Konráðs kom á markað var engin dönsk-íslensk orðabók til og því hefur hún verið keypt og notuð. Aðferð hans við að þýða oft fremur en nota jafnheiti hefur vafalaust valdið mörgum notandanum vonbrigðum sem var í leit að jafnheiti yfir danska flettiorðið. Sjálfur taldi Konráð það einn helsta kost bókarinnar að hann vísaði til fornrita. Með því fjarlægðist hann samt samtímamálið sem gerði það að verkum að bókin er blanda af gömlu og nýju og því ekki eins notendavæn og hún hefði getað verið. Í 3.4 er vísað til ummæla Björns Jónssonar um að bókin hafi þótt óhandhæg og ekki við alþýðu hæfi og er sá dómur Björns líklega nokkuð góð lýsing frá sjónarhóli hins almenna notanda tvímála orðabókar.

Heimildir

- An Icelandic-English dictionary / based on the MS. collections of Richard Cleasby, enlarged and completed by Gudbrand Vigfusson ; with an introduction and life of Richard Cleasby by George Webbe Dasent.* 1874. Oxford: Clarendon Press.
- An Icelandic-English dictionary / initiated by Richard Cleasby, subsequently revised, enlarged and completed by Gudbrand Vigfusson.* 2nd ed. with a supplement by Sir William A. Craigie. 1957. Oxford: Clarendon Press.
- Bjarni Vilhjálmsson. 1944. Nýyrði í Stjörnufræði Ursins. *Skírnir*, 118. árg.: 99–130. [Endurprentuð í *Orð eins og forðum*. Greinasafn eftir Bjarna Vilhjálmsson gefið út í tilefni sjötugsafmælis hans 12. júní 1985. Bls. 1–29. Reykjavík: Hafsteinn Guðmundsson.]
- Björn Halldórsson. 1814. *Lexicon Islandico-Latino-Danicum Björnnonis Haldorsonii*. Björn Haldorsens islandske Lexikon. Ex Manuscriptis Legati Arna-Magnæani cura R. K. Raskii editum. I–II. Havniæ: apud J. H. Schubothum.
- Björn M. Ólsen. 1891. Konráð Gíslason. *Tímarit Hins íslenzka bókmenntafjells* 12:1–96.
- Bréf = *Bréf Konráðs Gíslasonar*. 1984. Aðalgeir Kristjánsson bjó til prentunar. Reykjavík: Stofnun Árna Magnússonar.
- Briefwechsel = *Briefwechsel der Gebrüder Grimm mit nordischen Gelehrten*. 1885. Ernst Schmidt (Hrsg.) Berlin: Dümmler.
- Dansk Ordbog udgiven under Videnskabernes Selskabs Bestyrelse*. 1793–1848. I–VI. Kiøbenhavn.
- Eiríkur Jónsson. 1863. *Oldnordisk Ordbog ved det kongelige nordiske Oldskrift-Selskab*. Kjöbenhavn.
- Finnbogi Guðmundsson. 1970. Frá Hallgrími Scheving. *Landsbókasafn Íslands. Árbók 1969*, 26. ár, bls. 156–209.

- Fritzner, Johan. 1867. *Ordbog over det gamle norske Sprog*. Kristiania: Trykt paa Feilberg & Landmarks Forlag.
- Fritzner, Johan. 1886–1896. *Ordbog over det gamle norske Sprog*. I–III. Omarbeidet, forøget og forbedret Udgave. Kristiania: Den norske Forlagsforening.
- Gísli Brynjólfsson. 1952. *Dagbók í Höfn*. Reykjavík: Heimskringla.
- Guðrún Kvaran. 1991. Konráð Gíslason. Málfræðingur og orðabókahöfundur. *Skagfirðingabók. Rit Sögufélags Skagfirðinga*. 20:47–70.
- Guðrún Kvaran. 2008. Hallgrímur Scheving og staðbundinn orðaforði. *Íslenskt mál og almenn málfræði* 30:153–178.
- Gunnlaugur Oddsson. 1819. *Orðabók, sem inniheldur flest fágæt, framandi og vandskilin ord, er verda fyrir i dönskum bókum*. Kaupmannahöfn.
- Gunnlaugur Oddsson. 1991. *Orðabók sem inniheldur flest fágæt, framandi og vandskilin ord, er verda fyrir i dönskum bókum*. Ný útgáfa með íslenskri orðaskrá. Jón Hilmar Jónsson sá um útgáfuna ásamt Þórdísi Úlfarsdóttur. Orðfræðirit fyrri alda I. Reykjavík: Orðabók Háskólans.
- Jakob Benediktsson. 1969. Íslensk orðabókarstörf á 19. öld. *Andvari*. Nýr flokkur. Bls. 96–108.
- Jón Ingjaldsson. 1855. *Boðsbréf*. 28. marts 1855. Húsavík.
- Jónas Hallgrímsson. 1836. Af eðlisháttum fiskanna. *Fjölur. Ár-rit handa Íslendingum*. Annað ár, 3–14.
- Jónas Hallgrímsson. 1989. *Ritverk Jónasar Hallgrímssonar*. I–IV. Reykjavík: Svart á hvítu.
- Jónas Jónasson. Eiginhandarrit af danskri orðabók. Án titilblaðs. Í eigu Amtsbókasafnsins á Akureyri. [Ljósrit í eigu Guðrúnar Kvaran].
- Jónas Jónasson. 1896. *Ný dönsk orðabók með íslenskum þýðingum*. Reykjavík: Ísafoldarprentsmiðja.
- Jónas Kristjánsson. 1986. Íslensk sprogpolitik i 1800-tallet. *De nordiske skriftspråkenes utvikling på 1800-tallet* 3:134–147. Nordisk språksekretariats rapporter 7. Oslo.
- Kjartan Ottósson. 1990. *Íslensk málhreinsum*. Rit Íslenskrar málnefndar 6. Reykjavík: Íslensk málnefnd.
- Konráð Gíslason. 1847. Afrit af bréfi Konráðs til konungs er varðveitt á Stofnun Árna Magnússonar í Kaupmannahöfn (Árnasafn, KG, 31 a).
- Konráð Gíslason. 1851. *Dönsk orðabók með íslenskum þýðingum*. Kaupmannahöfn.
- Meyer, L. 1844. *Fremmedord-Bog, eller kortfattat Lexikon over fremmede, i det danske Skrift- og Omgangs-sprog forekommende Ord, Konstudtryk og Talemaader; tilligemed de i danske Skrifter mest brugelige, fremmede Ordforkortelser*. Anden, forøgede og forbedrede Udgave. Kjøbenhavn: Forlagt af J.H. Schuboths Boghandling.
- Molbech, Christian. 1833. *Dansk Ordbog indeholdende det danske Sprogs Stammeord, tilligemed afledede og sammensatte Ord, efter den nuværende Sprogbrug forklarede i deres forskiellige Betydninger, og ved Talemaader og Exempler oplyste*. Kiöbenhavn: Paa den Gyldendalske Boghandlings Forlag.

- Möbius, Theodor. 1866. *Altnordisches Glossar. Wörterbuch zu einer Auswahl alt-isländischer und alt-norwegischer Prosatexte*. Leipzig.
- Nordisk leksikografisk ordbog*. 1997. Henning Bergenholtz o.fl. Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 4. Oslo.
- Páll Eggert Ólason. 1950. *Íslenzkar æviskrár*. III. bindi. J–N. Reykjavík: Hið íslenska bókmenntafélag.
- Ritmálssafn Orðabókar Háskólans*: http://arnastofnun.is/page/gagnasofn_ritmalsafn.
- Sveinbjörn Egilsson. 1860. *Lexicon poëticum / conscripsit Sveinbjörn Egilsson; edidit Societas regia antiquariorum septentrionalium*. Hafniae.
- Ursin, G.F. 1842. *Stjörnufræði, ljett og handa alþíðu*. Videiar klaustri.

Lykilorð

orðabækur, orðabókafræði, orðaforði, málhreinsun

Keywords

dictionaries, lexicography, vocabulary, linguistic purism

Abstract

The article is concerned with the lexicographical work of Konráð Gíslason (1808–1891) who was a professor of Nordic philology at the University of Copenhagen and better known as a grammarian. After a short introduction I discuss *An Icelandic-English Dictionary* (1874), which Gíslason worked on for fourteen years in company with other Icelanders but had not the opportunity to prepare for printing. Next I turn to the *Dönsk-íslensk orðabók* (Danish-Icelandic dictionary) that Gíslason published in 1851. The main model for this work was *Dansk Ordbog* (Danish dictionary) by Christian Molbech, published in 1833. By comparing one column in Gíslason's dictionary with the corresponding one in Molbech's work it becomes evident that Gíslason used Molbech's text and examples extensively as a source. I then compare the words Gíslason regarded as loan words with Ludvig Meyer's book on loan words from 1844 since Molbech avoided such words in his dictionary. Other sources are also discussed, in particular it is investigated which of the numerous neologisms of the poet Jónas Hallgrímsson were taken up by Gíslason. Separate sections are devoted to the reception of the Danish-Icelandic dictionary by Gíslason's contemporaries, and the Danish-Icelandic dictionary by Jónas Jónasson that was published in 1896 is discussed as well as the reason why Gíslason's dictionary was not revised and published again. Chapter four deals with dictionaries that Gíslason worked on during various periods of his professional life without seeing any of them published.

Guðrún Kvaran
Stofnun Árna Magnússonar í íslenskum fræðum / Háskóli Íslands
Neshagi 16
107 Reykjavík
gkvaran@hi.is



Susanne M. Arthur

Are Oranges Yellow? *Appelsínugulur* as a Basic Color Term in Icelandic

1 Introduction

In their groundbreaking *Basic Color Terms* (1969), Brent Berlin and Paul Kay argue that basic color terms enter languages in a specific order in an evolutionary sequence. According to Berlin and Kay, a basic color term must fulfill four criteria (Table 1).

	A BASIC COLOR TERM MUST BE	EXPLANATION
I.	mono-lexemic	The meaning of the color term is not predictable from the meaning of its parts. This excludes for example terms such as 'yellowish,' 'reddish-brown,' or 'lemon-colored.'
II.	exclusive	The meaning of the term cannot be included in that of any other color term. This excludes for example terms such as 'crimson' or 'scarlet,' which are both terms for 'red.'
III.	unrestricted	A color term is not considered basic, if its use is restricted to a specific, narrow class of objects. This excludes for example terms that are only used for the color of hair, such as 'blond.'
IV.	psychologically salient	The term must be in frequent use and psychologically recognizable for a large number of people. This excludes for example rare or personally-defined color terms, such as the "the color of the rust on my aunt's Chevrolet" (Berlin and Kay 1969:6).

Table 1. *The four criteria of a basic color term, according to Berlin and Kay (1969).*

Orð og tunga 15 (2013), 121–139. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

Berlin and Kay (1969:2) determined the existence of eleven basic color terms that match all of the four characteristics: white, black, red, green, yellow, blue, brown, pink, purple, orange, and grey. Their conclusions for the establishment of an evolutionary sequence and the different language stages are summarized in Table 2.

I		II		III		IV		V		VI		VII
white	→	red	↗	green	→	yellow	↘	blue	→	brown	→	purple
black			↘	yellow	→	green	↗					pink
												orange
												grey

Table 2. *The introduction of basic color terms into a language, according to Berlin and Kay (1969).*

According to Berlin and Kay (1969), all languages can be assigned to one – and only one – of the seven stages at a certain point in a language’s development.¹ Concerning stage VII, Berlin and Kay (1969:21–22) note that “the remaining basic categories, purple, pink, orange, and grey, are quickly added to the lexicon and ... in no particular order.” They further explain that “[s]tage VII systems include all eight-, nine-, ten-, and eleven-term systems” (1969:22).

Kirsten Wolf (2006:189; 2009:223; 2010:110) maintains that Old Norse-Icelandic is an early stage VII language with eight basic color terms: *svaþr*, *hvítr*, *rauðr*, *grænn*, *gulr*, *blár*, *brúnn*, and *grár*. She notes (2006:187; see also 2010:110) that a ninth basic color term, *bleikr* (for pink), has been added to the color lexicon in Modern Icelandic, which in Old Norse-Icelandic denoted simply a light color or light shade of color as in “pale,” “wan,” “bleached,” “blond,” “fair,” or “light-colored.”² In accordance with Berlin and Kay’s definition, Modern Icelandic would, therefore, be a stage VII language with nine basic color terms, lacking basic color terms for purple and orange (see also Poisson 2011:22, 26).³ Jackson Crawford (2011), however, argues convincingly that the Icelandic *fjólublár* ‘blue like a violet,’ commonly used for purple, should be considered a basic color term despite its compound

1 Berlin and Kay’s study and evolutionary sequence of color-terms are certainly not undisputed and have been revised since the publication of *Basic Color Terms* (1969). For critiques, see e.g., Hickerson 1971; Witkowski and Brown 1977.

2 For a more detailed discussion of *bleikr* in Old Norse-Icelandic, see Wolf 2005.

3 Wolf (2010:110) mentions that “*lilla* (purple) ... has very recently entered the language, and *órans* (orange) is in the process of entering the language.” Jackson Crawford and Susanne Arthur’s (née Fahn) research on these two color terms, however, has since shown that this statement may have to be revised (see below for *órans* and Crawford 2011 for *lilla*).

nature.⁴ Crawford (2011) insists that “the value of [Berlin and Kay’s] monolexemic criterion [– which excludes compound terms –] may be questioned in a language community that strongly identifies itself with independence from lexical borrowings from other languages.” As in the case of purple, Modern Icelandic uses a compound, *appelsínugulur* ‘yellow like an orange’ to denote the color orange (a secondary color-term according to Berlin and Kay’s study). This article seeks out to prove, however, that *appelsínugulur* – like *fjólublár* – should be considered a basic color term. The author begins with a brief history of the fruit and color orange (ch. 2) and gives an overview of the development of terms for the fruit and color in Icelandic (ch. 3 and ch. 4). She then explores reasons why *appelsínugulur* became the predominant term in Icelandic and examines whether it is perceived as a shade of yellow or an independent and basic color term (ch. 5 and ch. 6).

2 A Brief History of the Fruit and Color Orange

The orange tree probably originated in northern India and reached China around 2200 B.C. and later Europe via the Silk Road (Toussaint-Samat 1992:662–663; Root 1980:303). The term orange for the fruit is possibly of Dravidian (Tamil) origin and was adopted into Sanskrit (*nāraṅgaḥ*), and later Persian (*nārang*), Arabic (*nāranj*) and Spanish (*naranja*) (Barnhart 1988:733; OED 1989:X.887; OED online 2011: ‘orange;’ American Heritage 2009). The first oranges introduced to Europe – so-called Persian oranges – were bitter. The Latin term for this bitter variety is *pomum aurantium* ‘orange apple,’ from which, for example, the Italian *pomeransi*, French *pume orange*, German *Pomeranze*, and Danish *pomerans* are derived. *Aurantium* likely refers to the original Sanskrit term for the fruit, although it has been argued (see e.g., Bonavia 1888:327) that it is related to *aurum* ‘gold,’ thus referring to the golden color of orange peel.⁵ The English term *orange* is a shortening of the Old French form *pume orange*. The first recorded written citation of the word in Middle English is found in a text composed around 1380 (Barnhart 1988:733). However, a record from around 1290 suggests that oranges

⁴ I would like to thank Jackson Crawford for providing me with a copy of his unpublished conference paper.

⁵ Another Latin term used to describe oranges is *malum aureum* ‘golden apple’ (Bonavia 1888:327).

had already reached England during that time. This document indicates that queen Eleanor of Castille bought seven oranges after a cargo ship from Spain arrived in Southampton (Root 1980:304).

The bitter Persian oranges were replaced in the fifteenth century by a sweet variety, brought to Europe from India by Portuguese traders (Harper 2001). This sweet variety reached northern Germany by the beginning of the eighteenth century, via Amsterdam and Hamburg (Drosdowski 1989:40). The German term *Apfelsine* (derived from Old Dutch: *appelsina*, Low German: *Appelsina*) for the sweet orange is a literal translation of 'China-apple'⁶ from an old word *Sina* for China, because oranges were brought to Europe from South China around the year 1500 (Drosdowski 1989:40). In Germany, the term *Orange* for the fruit was generally used in the south, while *Apfelsine* was more prevalent north of the river Main (Ammon et al. 2004:545), indicating that the term 'China-apple,' which is also used in all Scandinavian languages (Danish: *appelsin*, Norwegian-Bokmål: *appelsin*, Norwegian-Nynorsk: *appelsin*, Swedish: *apelsin*, Icelandic: *appelsína*, Faroese: *appilsin*), is typical for Northern Europe. In Denmark and Sweden, references to oranges can be found in documents from the sixteenth and seventeenth centuries (Heizmann 2005:43). It is unclear when oranges were first introduced to Iceland (see ch. 3).

The color term orange, which is derived from the name of the fruit and likely a shortening of 'orange-colored,' was seemingly not used in English before 1542 (Ammon et al. 2004:545; *Ordbog over det Danske Sprog* 1975:XV.1174; Harper 2001). Despite the fact that the fruit is generally called 'China-apple' in Scandinavia, most Scandinavian languages use a word derived from orange to indicate color (Danish: *orange*, Norwegian-Bokmål: *oransje*, Norwegian-Nynorsk: *oransje*, Swedish: *orange*). The only exceptions are Icelandic and Faroese, which have developed compound terms or secondary color terms according to Berlin and Kay (Icelandic: most commonly *appelsínugulur*, Faroese: *brandgulur*, *reyðgulur*).

3 The Fruit Orange in Icelandic

Icelanders have used a number of terms to describe the orange fruit (Table 5), which are largely derived or translated from Latin (*pomum*

⁶ In the eighteenth century the term *Chinaapfel* was also used in Germany in addition to *Apfelsine* (Drosdowski 1989:40).

aurantium, malum aureum), English (*orange*), German (*Pomeranze, Apfelsine, Orange*), Danish (*appelsin, pomerans*) or other Scandinavian forms (*appelsin/apelsin*). The oldest Icelandic term *eyjarepli* occurs only once, in *Kirialax saga*,⁷ a saga whose composition is dated to the mid-fourteenth century, but it is doubtful that the term in fact refers to oranges or that the fruit was known in Iceland at that time (Heizmann 2005:42–43, 47–48).⁸

Mention of the fruit appears in the mid-seventeenth century and becomes more frequent from the first half of the eighteenth century onwards. The earliest examples – listed in Table 3 – are derived from the Latin *pomum aurantium*, German *Pomeranze* and Danish *pomerans*. Quite often the term appears in texts translated from German or Danish (see Table 3).

In three cases (Gudmund Thorgrimsen 1784:25; Magnús Stephensen 1797:266; Steingrímur Thorsteinsson 1886:199–200, 229), *pomerans* is used parallel to the Icelandic term *gullepli* ‘golden apple’ (see Table 3 and Table 4), likely a translation of the Latin *malum aureum* (see fn. 5). *Ritmálssafn Orðabókar Háskólans (ROH)* lists seventeen occurrences of *gullepli* from the mid-eighteenth century to the late twentieth century. The term is problematic, since it may not necessarily just refer to the fruit orange, but could also describe an apple made of gold. The oldest example is in Jón Jónsson’s *Meditationes triumphales. Edur Sigurhrooss Hugvekiur* (1749). It is doubtful that the term refers to an orange in this case, since the phrase reads “Þu ert O IESu sem eitt Gull-Eple i Silfur-Skaal” (Jón Jónsson 1749:95) [‘You are, oh Jesus, like a golden apple in a silver bowl’];⁹ here the author is clearly contrasting the gold of the apple with the silver of the bowl. The second oldest example of *gullepli* (Gudmund Thorgrimsen 1784:25), on the other

7 “hennar höfuð var því líkt sem eyjar epli með gulligum lokkum” (Konráð Gíslason 1860:406; Kálund 1917:75–76) [Her head was like an ‘island apple’ with golden curls.] – The quote is from the description of Florencias, the daughter of king Lotharius of Byzantine and includes comparisons to plants and minerals. See also Heizmann (2005:42).

8 Kálund notes in his edition of *Kirialax saga* that the term is a shortening of *Sikileyiar epli* ‘apple from Sicily’ and means *appelsin* (1917:76). His etymology has since been adopted (see for example *Íslensk orðabók* (ÍO 2007:196) ‘*eyjarepli*,’ or *Wikipedia. Frjálsa alfræðiritið: ‘appelsína.’*). In her English translation of *Kirialax saga*, Alenka Divjak (2009:338) translates the term with ‘orange,’ without suggesting that the term might be questionable. Heizmann (2005), however, convincingly challenges Kálund’s assumption.

9 See Vulgate, Prov. 25.11 “Mala aurea in lectis argenteis.”

hand, clearly refers to the fruit orange (see Table 3).¹⁰ In the late nineteenth and twentieth centuries, the term becomes increasingly used in poetry rather than prose (see *ROH* for examples), and in two cases it occurs in a glossary (Steingrímur Thorsteinsson 1886:199, 229), next to other possible terms for the orange fruit (see Table 3 and Table 4).

QUOTE	DATE/SOURCE	TRANSLATED FROM
"En engin aldini voru á þeim tíma árs fullvaxin utan <i>pómerans</i> og <i>bónanster</i> "	1661; Guðbrandur Jónsson 1946: II.73	
"margslags Triaam ... <i>Pomerans</i> og <i>Laurberia Triaam</i> "	Lassenius 1723:150	German; Lassenius 1857:80–81
"Paalmavidar-Tried merker hans Sigurvinning ... enn <i>Pomerants</i> Avøxturenn þann sæta Ylm ..."	Lassenius 1723:151	German; Lassenius 1857:80–81
"I Gardenum feck hann mœrg aagiæt Aldine, so sem ... <i>Pomerantz-Eple</i> "	Björn Markússon 1756:314	Danish
"Aldinin eda Eplin. ... Gull-epli (mala aurea, <i>Pomerantzer</i>)"	Gudmund Thorgrimsen 1784:25	German; Büsching 1779:80
"Eyian er ynnðæl, øll yrke, pløntud og þakin granat-epla-, fikiu- og gull-epla edur <i>pómerants-tríam</i> "	Magnús Stephensen 1797:266	
"I landi þessu finna menn ... sudurlanda ávexti (<i>Pómeranzur</i> , ...)"	Grímur Jónsson 1821:II.1, 224	
"Gold...–orange, kvk. ... gullepli, <i>pómeranza</i> (pomum aurantium), <i>óransía</i> ."	Steingrímur Thorsteinsson 1886:199–200	German (glossary)

Table 3. Earliest examples of terms derived from the Latin pomum aurantium, German Pomeranze or Danish pomerans.

By the middle of the nineteenth century, when, presumably, oranges were still a fairly new and exotic product in Iceland and no common Icelandic term for the fruit existed, the terms *orangeávöxtur* 'orange-fruit,' *orange-epli* 'orange-apple,' and *óransía* 'orange' appear in written sources. *ROH* only lists one example for each term (Table 4).

QUOTE	SOURCE
" <i>Orange-ávextir</i> (appelsínur) voru fluttir frá Kína til Portúgal, 1547, og komust þaðan út um suðurhluta Evrópu"	Páll Sveinsson 1862:94
"Ein tegundin verðr eins stór og smá <i>orange-epli</i> "	Jón Jónsson Hjaltalín 1862:33
"Orange, ... kvk. ... <i>óransía</i> , <i>óransíutré</i> , gullepli, gullapaldur."	Steingrímur Thorsteinsson 1886: 229

Table 4. Earliest examples of terms derived from orange.¹¹

¹⁰ Gudmund Thorgrimsen uses *gullepli* to translate the German term *Pomeranze*. See Büsching 1778:80.

¹¹ It is possible that *orange-epli* refers to the Cox Orange Pippin apple rather than an

Appelsína, the only used term in Modern Icelandic, appears in print in the middle of the nineteenth century in a recipe for “*Rúgbrauðs kaka*” ‘rye bread cake’ in Þóra Andrea Nikólína Jónsdóttir’s *Ný matreiðslubók ásamt ávísun um litun, þvott o.fl.* (1858:119). The term is a Danish or German loanword (see e.g. Ásgeir Blöndal Magnússon 1989:22). Considering that recipes are generally written to be intelligible to all, it is obvious that *appelsína* was the predominant term by the time the cookbook was printed. Similarly, Páll Sveinsson (1862:94) states “Orange-ávéxtir (*appelsínur*) voru fluttir frá Kína til Portúgal, 1547, og komust þaðan út um suðurhluta Evrópu” [‘Orange-fruits (China-apples) were brought from China to Portugal in 1547, and spread from there through the southern parts of Europe’]. The quotation shows that *appelsína* was in this case used simultaneously alongside *orange-ávéxtir* and was presumably equally or more readily understood at the time.

Attempts to introduce the Icelandic term *glóaldin* ‘glowing fruit’¹² in the early twentieth century to replace the loanword *appelsína* failed. This was presumably because, as Jón Hilmar Jónsson (1978:358) suggests, the loanword entered the language quickly and was widely accepted in a short time. *ROH* lists thirteen citations of *glóaldin* dating between 1929 (Helgi Pjeturss 1929:228) to 1972 (Halldór Laxness 1972:237).

TERM	EARLIEST OCCURRENCE
eyjarepli ?	mid-14th century
terms derived from <i>pomerans</i>	1723
gullepli	(1745) 1784
terms derived from <i>orange</i>	1862
<i>appelsína</i>	1858
<i>glóaldin</i>	1929

Table 5. Summary of Icelandic terms for the fruit orange (order as discussed in article).

actual orange. This kind of apple is available in Iceland, as for example illustrated by newspaper advertisements, such as in *Þjóðviljinn* (1951:7) or *Dagur* (1960:6), where *Cox Orange epli* ‘Cox Orange apples’ are advertised next to *appelsínur* ‘oranges.’ It should be noted, however, that these advertisements are much younger than Jón Jónsson Hjaltalín’s article. It is, therefore, uncertain if Cox Orange apples were known in the nineteenth century. Jón Jónsson Hjaltalín (1862:33) uses the term as a size comparison in an article about echinoderms, which means that the context does not allow for a definite interpretation of the term.

12 In the author’s opinion, *glóaldin* may be a loose translation of the Latin *malum aureum*.

4 The Color Orange in Icelandic

Before the introduction of the fruit orange and, consequently, the term *appelsínugulur*, Icelanders appear to have used a variety of terms to describe orange-colored objects, largely compounds of red and yellow.

Ordbog over det norrøne prosasprog (ONP) lists five examples of *rauðbleikr*. Kirsten Wolf (2010:123) suggests that *bleikr* was used to describe the hue yellow before the introduction of the term *gulr*. Accordingly, *rauðbleikr* could be considered a compound of red and yellow. The citations in *ONP*, dating from the late thirteenth to the late fourteenth century,¹³ are used to refer to hair, a beard, and two gemstones (sardonyx and jacinth), all of which may be perceived as orange-colored. While Cleasby and Vigfússon's translation of *rauðbleikr* as 'reddish' (1874:484) is acceptable, it conceivably refers to a color in the range between red and yellow.

The most frequently used term for the color orange in Icelandic (besides the now prevalent term *appelsínugulur*) is *rauðgulur* 'red-yellow,' which appears three times in *ONP* and fifty-three times in *ROH*. The oldest occurrence in *ONP* is from a testament written in Oslo on April 13, 1331 (DN 1855:III.148).¹⁴ It refers to a *pallklæde* 'carpet/covering' and could describe a solid orange color or a red and yellow pattern. The other occurrences are in *Sturlunga saga* (AM 122 a fol., c1350–1370) and *Karlamagnús saga* (AM 180 d fol., c1700),¹⁵ respectively, and both describe the color of hair (*ONP*). As with *rauðbleikr*, these instances describe objects that can be conceived of as orange-colored. *ROH*, which contains examples from after the middle of the sixteenth century, lists occurrences of *rauðgulur* dating from 1545 to 1977. Until the end of the seventeenth century, the term was almost exclusively used to denote the color of hair, clothing and fabrics, particularly altar linens. The adjective was later associated with a larger variety of objects, although its use for ginger-colored hair still prevailed (see *ROH*

13 One of the examples is found in AM 447 4to, a late seventeenth-century manuscript containing *Eyrbyggja saga* (see *ONP*). The occurrence of *rauðbleikr* in this case cannot be dated to the thirteenth to fourteenth century with certainty since it could be a post-medieval variant.

14 Since the example is taken from a Norwegian source, it cannot be included in this study, which focuses on Icelandic. It is nonetheless mentioned for completeness.

15 As with AM 447 4to (see fn. 13), the occurrence of *rauðgulur* in this case cannot be dated to the Middle Ages with certainty because the citation exists only in a post-medieval manuscript (AM 180 d fol.).

for examples). In the glossary of *Þýzk lestrarbók* (Steingrímur Thorsteinsson 1886:229) the German color *orange* is defined as something “sem hefir óransíu lit, eða er rauðgulur eins og óransíubörkur” [that is the color of an orange, or red-yellow like the rind of an orange]. Here, the author makes a definite connection between *rauðgulur* and the fruit and color orange.

From the mid-seventeenth century (1642) to the late twentieth century (1992) the term *brandgulur* ‘fire-yellow’ appears. The thirteen examples in *ROH* suggest that the term is closely associated with textiles – especially silk – and in the twentieth century with the color of mushrooms and animals. Just as *bleikr* was used to indicate the hue yellow, *brandr* ‘brand/flame’ – referring to fire – may have been used to indicate the hue red.¹⁶ This supports the argument that before the introduction of the fruit Icelandic only used compounds of red and yellow for the color orange.

Twelve examples of *gulrauður* ‘yellow-red’ dating between 1704 and 1977 are listed in *ROH*. The term is used in a variety of contexts, including the description of textiles. It does not, however, refer to ginger hair for which *rauðbleikr*, *rauðgulur*, and *gullrauður* (see below) appear to be the only compounds used.

In the late nineteenth century, isolated examples of *oransjegulur* (1877), *orangegulur* (1882), and *óransíugulur* (1891) – all meaning ‘orange-yellow’ – appear (*ROH*). This coincides neatly with the nineteenth-century examples of *orangeávöxtur* ‘orange-fruit,’ *orangeepli* ‘orange-apple,’ and *óransía* ‘orange’ noted above. Icelanders presumably noticed the close connection between the fruit and the color, which they may have encountered in other languages, but since no common Icelandic term for the fruit – and therefore the color – had yet been introduced into the Icelandic lexicon, a variety of terms emerged. *ROH* lists one example of *gulleplalitaður* ‘colored like a golden apple’¹⁷ from 1906 and nine examples of *gullrauður* ‘golden-red’ from between 1935 and 1977. The latter, however, may describe a ‘golden/shimmering’ red color more than the actual color orange.¹⁸ The term is frequently

16 Kirsten Wolf (2010:122) gives an example from *Skáldskaparmál*, in which gold is described as “fire of arm or joint or limb, since it is red.” This may suggest that both gold and fire were considered red in Old Norse-Icelandic.

17 A reference to the Latin *malum aureum* (ch. 2 and ch. 3).

18 As with *bleikr*, Kirsten Wolf (2010:123) suggests that *gull* may have been used to describe the hue yellow before the introduction of *gulur*. This suggests that *gullrauður*, like *rauðbleikr* could be considered a compound of yellow and red.

associated with the description of flames, the sun (sunrise/sunset), copper, and the color of hair (see *ROH* for examples).

The term *appelsínugulur* that is most commonly used now, first appeared during the middle of the twentieth century (see *ROH* for oldest examples), and, just like the loanword *appelsína*, it caught on much faster than the Icelandic creation *glóaldin*. In terms of frequency, the color term *appelsínugulur* outnumbers *glóaldinrauður* ‘red like an orange,’ for which *ROH* lists only one example from 1966. Occurrences of *appelsínurauður* ‘red like an orange’ also exist (*ROH*), but are less frequent than *appelsínugulur*.

In an informal survey taken by a small number (6) of Icelanders, participants were asked to list any other Icelandic color terms which they might use instead of *appelsínugulur* for the color orange.¹⁹ Two participants did not list any alternatives. One mentioned *rauðgulur* and “*órans*” ‘orange;’ another *órans*, *rauðgulur*, *gulrauður*, and *appelsínulitur* ‘orange-colored;’ the third *dökkgulur* ‘dark-yellow,’ *ljósrauður* ‘light-red,’ *sólarlag* ‘sundown,’ and *roði* ‘redness/blush;’ and the fourth *rauðgulur* and *appelsínurauður*.²⁰ The use of *órans* may suggest that Icelandic is beginning to incorporate a true basic color term – clearly a loan word – instead of a compound term for the color orange. The fact that one of the participants put the term into quotation marks, however, indicates that the term is likely considered uncommon.²¹ Since *appelsínugulur* is well established, it is doubtful that *órans* should become predominant.

19 The survey had four major parts. First, participants (of varying genders, ages and educational backgrounds) were shown pictures of different objects and color palettes. The objects and palettes were yellow, orange, red, or shades thereof. The participants were only allowed to describe the objects as ‘yellow/gulur’ or ‘red/rauður.’ In the second part, the same pictures were shown. This time participants should write down any Icelandic color term they would use to describe the objects in question. Next, participants were asked to indicate which squares of a detailed color palette they considered *orange* (English) and *appelsínugulur* (Icelandic). Lastly, participants answered the question whether they thought the fruit orange was named after the color or vice versa, and listed Icelandic color terms besides *appelsínugulur* which they might use for the English term *orange*. They were instructed to arrange their answers from ‘most likely to use’ to ‘least likely to use.’

20 Of these answers, *sólarlag* and *roði* should definitely be eliminated since they are not adjectives. *Appelsínulitur* can be a variant of the adjective *appelsínulitaður* ‘orange-colored’.

21 The term is not included in the *Íslensk orðabók* (ÍO 2007) or any online database. A larger survey (possibly online) would be needed to determine if *órans* is in fact entering the Icelandic lexicon. During time spent in Iceland, the author of this article has, however, never heard or seen the term *órans* used in conversation or writing, which suggests that such a survey may not be fruitful or valuable.

TERM	EARLIEST OCCURRENCE
rauðbleikr	late 13th century
rauðgulur	14th century
brandgulur	mid-17th century
gulrauður	early 18th century
oransjegulur, orangegulur, óransíugulur	late 19th century
gulleplalitaður	early 20th century
gullrauður	mid-20th century
appelsínugulur	mid-20th century
glóaldinrauður	1966
appelsínurauður	mid-20th century
appelsínulitur/appelsínulitaður	21st century?
dökkgulur	21st century?
ljósrauður	21st century?

Table 6. Summary of Icelandic compound terms for the color orange (order as discussed in article; the final three examples appeared only in the survey).

5 Assessment of Data: The Story of the Yellow Orange

Of the thirteen examples of Icelandic compound color terms for orange (Table 6), five are definite or possible compounds of red and yellow (*rauðbleikr*, *rauðgulur*, *brandgulur*, *gulrauður*, *gullrauður*), three are associated only with yellow (*oransjegulur/orangegulur/óransíugulur*, *appelsínugulur*, *dökkgulur*), three are associated only with red (*glóaldinrauður*, *appelsínurauður*, *ljósrauður*), and two (*gulleplalitaður* *appelsínulitur/appelsínulitaður*) resemble 'orange-colored,' the word from which orange is ultimately derived.

It is to be expected that orange is frequently described as a compound of red and yellow, or becomes isolated from either red or yellow, in early language stages, since it is not a primary color but is formed by mixing those two colors. In Berlin and Kay's analysis of early language stages (stages I–III), which do not yet have a term for yellow, both yellow and orange often seem to derive from red (1969:26, 29). After the introduction of yellow into a language – occasionally as a loanword from other languages – the color orange, however, is frequently associated with yellow.²² Considering the close

²² Examples include the language of the !Kung Bushmen in South Africa (Berlin and Kay 1969:33), certain languages of the Torres Straits groups (Berlin and Kay 1969:38–39), the language of the Native American Zuni (Lenneberg and Roberts

connection between the colors orange and yellow, it is not surprising that Berlin and Kay (1969:22) suggest that, when orange is established as a separate basic color term in a language, it most often “becomes isolated from YELLOW.” There are, however, other instances where orange is derived from red or considered part of red.²³

The examples of compound color terms for orange in Modern Icelandic, support the view that orange derives from either red or yellow. The oldest color terms describing orange objects (*rauðbleikr*, *rauðgulur*, *brandgulur*) as well as the terms coined when oranges became better known in Iceland (*oransjegulur*, *orangegulur*, *óransíugulur*) use a hue of yellow (*-bleikr/-gulur*) as a suffix, indicating that orange may originally have been perceived as a shade of yellow by Icelanders.²⁴ Due to the strong historical association with yellow, a direct translation of ‘orange-colored’ (*gulleplalitaður*, *appelsínulitaður*) or a term with the suffix *-rauður* would not have been adequate after the introduction of the fruit to Iceland. This may explain why *appelsínugulur* became the predominant term. The prefix *appelsínu-* was likely given preference over *rauð-*, *brand-*, or *dökk-* because of the close connection between fruit and color. Since the loanword *appelsína* entered the Icelandic lexicon very quickly, it likely influenced the creation of the new color term.

6 *Appelsínugulur* as a Basic Color Term

Though orange may have originally been perceived as a hue of yellow in Iceland, it can be argued that today, *appelsínugulur* should be considered an independent and basic color term despite its compound nature. In a recent study, Regier, Kay and Khetarpal (2007) argue that “Color Naming Reflects Optimal Partitions of Color Space.”

1956:31), or certain dialects in southwest Iran, where orange can be included in yellow or red, depending on the shade, but it is most frequently associated with yellow (Friedl 1979:58, fig. 6; 63, fig. 9; 64). Furthermore, Delwin and Brown (2006:16608–9) illustrate in an analysis of the “World Color Survey (WCS) color-naming data set by using k-means cluster and concordance analyses” that yellow and orange are inseparably connected by establishing a WCS chromatic category “YELLOW-OR-ORANGE.”

²³ Examples include the language Malaya (Berlin and Kay 1969:97), the language of the Aguaruna, the indigenous people of the Peruvian jungle (Berlin and Berlin 1975:68), or Jaqaru, a language spoken in the Andes Mountains in Peru (Hardman 1981:66).

²⁴ The fact that orange juice generally has a yellowish color may add to this perception, since there is such a close connection between the fruit and color orange.

They nonetheless admit (Regier, Kay and Khetarpal 2007:1441) that “linguistic convention may be one force that can pull a particular language away from a perceptually optimal partitioning of color space” and that the historical development of a language must be taken into consideration. Both the history of the language and linguistic convention play an important role in Iceland. Icelandic grammar and language purism favor the creation of new compound words rather than the adaptation of monolexemic loanwords, which makes the monolexemic criterion of basic color terms unsuitable for such a language community.²⁵ Berlin and Kay’s monolexemic criterion can also be criticized. While they consider *orange* to be a basic color term (fulfilling the monolexemic criterion), they eliminate the examples “lemon-colored” and “salmon-colored” for not being monolexemic (1969:5–6). Since orange is, however, a shortening of orange-colored, its status as a basic color term could conceivably be questioned as well.

Appelsínugulur fulfills Berlin and Kay’s remaining three criteria. According to the two scholars (1969:6), a basic color term must apply to a wide range of objects. *Appelsínugulur* is used unrestrictedly to denote orange-colored objects.²⁶ Moreover, *appelsínugulur* fulfills the fourth criterion as it is “psychologically salient for informants” (Berlin and Kay 1969:6). It is the predominant term for orange. In fact, in the *Íslensk-ensk orðabók* (Sverrir Hólmarsson, Christopher Sanders and John Tucker 1989) it is the only term listed for the color orange. Its preferential status over other possible color terms is also supported by the fact that two survey participants did not list any alternatives to *appelsínugulur*. Additionally, in the survey, Icelanders almost always uniformly used *appelsínugulur* (occasionally with the common prefixes *dökk-* ‘dark,’ *ljós-* ‘light,’ and *skær-* ‘bright’) for describing orange objects. When asked to indicate which numbered color squares in a color palette to be *appelsínugulur* (Icelandic) or *orange* (English), the participants listed the exact same number range for both languages, which reveals that all participants considered Icelandic *appelsínugulur* and English *orange* as equivalent. A similar informal survey conducted by the author among native German-speakers, who use basic color

25 Jackson Crawford (2011) makes the same argument with regards to *fjólublár* ‘purple.’ See fn. 5.

26 Ásta Svavarsdóttir, editor of *Orð og Tunga*, brought to my attention that *appelsínugulur* would, however, never be used for hair in the Modern Icelandic, for which *rauður* is chosen. This is comparable to English, where a reddish color of hair is generally referred to as *red* or *ginger* rather than *orange*.

terms for yellow (*gelb*) and orange (*orange*), resulted in comparable number ranges, particularly with regard to the average starting point in the red area of the spectrum. There is no doubt that there is a uniform consent amongst Icelanders of what is conceived of as *appelsínugulur* and that the term is distinct from yellow.

One survey participant described a dark-orange object as *rauð-appelsínugulur* ‘red like a yellow orange’ and, as noted, some participants occasionally indicated nuances of orange with the common prefixes *dökk-*, *ljós-*, and *skær-*. The fact that shades of orange are described by adding prefixes to the term *appelsínugulur* (rather than *gulur* or *rauður*), supports the argument that *appelsínugulur* is its own lexemic unit, clearly defined and distinct from *gulur* and *rauður*. It would not be necessary to add the prefix *rauð-* to a compound color term (*appelsínugulur*), if the compound itself was considered a subclass of another color (*gulur*).²⁷ There are multiple examples of *rauðgulur*. The term can undoubtedly be used to describe dark orange objects, rendering a creation such as *rauðappelsínugulur* unnecessary. The only explanation appears to be that *appelsínugulur* is considered a lexemic entity in itself rather than a secondary color term, a shade or subclass of yellow. Accordingly, it fits Berlin and Kay’s second criterion for a basic color term in that “it is not included in that of any other color term” (1969:6).

7 Conclusion

The examples of color terms for orange in Modern Icelandic and the survey results reveal that in the past, Icelanders clearly associated the color with yellow more than with red. While the dominant term *appelsínugulur* may originally have been perceived as a shade of yellow, this study shows that it is now psychologically salient and not necessarily considered a sub-class of yellow. *Appelsínugulur* should be considered a basic color term despite its compound nature, as the monolexemic criterion put forth by Berlin and Kay is problematic. Correspondingly, Modern Icelandic should be considered a Stage VII language with at least ten, if not all eleven basic color terms: *svartur*,

²⁷ Comparably, while descriptive color terms such as ‘reddish yellow,’ ‘reddish orange,’ ‘orange-yellow,’ or ‘orange-red’ are easily acceptable and understandable in English, a color term such as ‘reddish orange-yellow,’ consisting of three individual color terms, would be very uncommon.

hvítur, rauður, grænn, gulur, blár, brúnn, grár, bleikur, (likely) *fjólublár* (see Crawford 2011), and (most certainly) *appelsínugulur*.

Bibliography

- American Heritage = “orange.” *The American Heritage® Dictionary of the English Language*. 2007. Boston: Houghton Mifflin. On the Credo Reference website: <http://www.credoreference.com/entry/7109030> (October 31, 2011).
- Ammon, Ulrich, et al. 2004. *Variantenwörterbuch des Deutschen: Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol*. Berlin/New York: Walter de Gruyter.
- Ásgeir Blöndal Magnússon. 1989. *Íslensk orðsifjabók*. Reykjavík: Orðabók Háskólans.
- Barnhart, Robert K. (ed.). 1988. *The Barnhart Dictionary of Etymology*. [Bronx, N.Y.]: The H.W. Wilson Company.
- Berlin, Brent and Paul Kay. 1969. *Basic Color Terms*. Berkeley: University of California Press.
- Berlin, Brent and Elois Ann Berlin. 1975. Aguaruna Color Categories. *American Ethnologist* 2.1: 61–87.
- Björn Markússon. 1756. *Þess Svenska Gustav LandKRONNS Og Þess Engelska Bertholds Faabreitileger Robinsons, Edur Liifs Og Æfe Sögur*. Hólar: [n.n.]. On the Bækur.is website: <http://baekur.is/> (June 27, 2012).
- Bonavia, Emanuel. 1888. *The Cultivated Oranges and Lemons, Etc., of India and Ceylon*. London: W. H. Allen & Co.. On the Google Books website: <http://books.google.co.uk/> (June 27, 2012).
- Büsching, Anton Friedrich. 1778. *Unterricht in der Naturgeschichte*. (3rd edition). Berlin: C. J. A. Wangens Wittwe. On the Google Books website: <http://books.google.co.uk/> (June 27, 2012).
- Cleasby, Richard and Guðbrandur Vigfússon. 1874. *An Icelandic-English dictionary*. Oxford: Clarendon Press. On the Germanic Lexicon Project website: http://lexicon.ff.cuni.cz/texts/oi_cleasbyvigfusson_about.html (June 27, 2012).
- Crawford, Jackson. 2011. *Basic Color Terms in Icelandic: The Case of Fjólublár*. The Germanic Linguistics Annual Conference 17, Austin, Texas, 15–16 April 2011. [unpublished].
- Dagur* 57. December 12, 1960. On the Tímarit.is website: <http://timarit.is/> (June 27, 2012).
- Divjak, Alenka. 2009. *Studies in the Traditions of Kirialax Saga*. Ljubljana: Inštitut Nove revije, zavod za humanistiko.
- DN = *Diplomatarium Norvegicum*. 22 vols. 1848–1995. Christiania (Oslo): P. T. Mallings Forlagshandel.
- Drosdowski, Günther. 1989. *Duden Etymologie: Das Herkunftswörterbuch*. (2nd

- edition) Duden in 10 Bänden 7. Mannheim: Dudenverlag.
- Friedl, Erika. 1979. Colors and Culture Change in Southwest Iran. *Language in Society* 8.1: 51–68.
- Grímur Jónsson. 1821. *Almenn Landaskipunarfræði*. 2 vols. Copenhagen: Hið íslenska bókmenntafélag. On the Bækur.is website: <http://baekur.is/> (June 27, 2012).
- Guðbrandur Jónsson (ed.). 1946. *Reisubók Jóns Ólafssonar Indíafara. Samin af honum sjálfum 1661*. 2 vols. Reykjavík: Bókfallsútgáfan.
- Gudmund Thorgrimsen. 1784. Framhalld Doct. Anton Friderich Büschings Undirvisunar í Náttúru-Historiunni. *Rit þess (konunglega) íslenska Lærdómslistafélags* 5: 1–32. On the Tímarit.is website: <http://timarit.is/> (June 27, 2012).
- Halldór Laxness. 1972. *Guðsgjafapula*. Reykjavík: Helgafell.
- Hardman, Martha J. 1981. Jaqaru Color Terms. *International Journal of American Linguistics* 47.1: 66–68.
- Harper, Douglas. 2001–2011. *Online Etymology Dictionary*. On the website: <http://www.etymonline.com/> (June 27, 2012).
- Heizmann, Wilhelm. 2005. Orangen und Zitronen im alten Island? Das *eyjar-epli* der *Kirjalax saga*. In: Thomas Seiler (ed.), *Herzort Island. Aufsätze zur isländischen Literatur- und Kulturgeschichte. Festschrift für Gert Kreutzer*, pp. 42–48. Lüdenscheid: Seltmann.
- Helgi Pjeturss. 1929. *Ennýall: nokkur íslensk drög til skilnings á heimi og lífi*. Reykjavík: [n.n.].
- Hickerson, Nancy P. 1971. Review of Basic Color Terms: Their Universality and Evolution. By Brent Berlin and Paul Kay. Berkeley and Los Angeles: University of California Press, 1969, Pp. xi, 178. *International Journal of American Linguistics* 37.4: 257–270.
- ÍO = *Íslensk orðabók*. 2007. (4th edition) Editor: Mörður Árnason. Reykjavík: Edda.
- Jón Jónson Hjaltalín. 1862. Athugasemdir um sullaveikina á Íslandi. *Þjóðólfrur* 15.8–9: 33–34. On the Tímarit.is website: <http://timarit.is/> (June 27, 2012).
- Jón Jónsson. 1749. *Meditationes triumphales. Edur Sigurhrooss Hugvekiur*. Hólar: [n.n.]. On the Bækur.is website: <http://baekur.is/> (June 27, 2012).
- Jón Hilmar Jónsson. 1978. Zur Sprachpolitik und Sprachpflege in Island. *Muttersprache* 88: 353–362.
- Konráð Gíslason (ed.). 1860. *Fire og fyrretyve for en stor deel forhen utrykte prøver af oldnordisk sprog*. Copenhagen: Gyldendalske Boghandling F. Hegel. On the Google Books website: <http://books.google.co.uk/> (June 27, 2012).
- Kálund, Kristian (ed.). 1917. *Kirialax saga*. Copenhagen: S. L. Møller.
- Lassenius, Johann. 1723. *Tvennar sjö sinnum sjö hugvekiur eða þankar út af písl og þínu Drottins vors Jesú Kristi, sem lesast mega á kvöldin og morgnana, um allan föstutímamann*. Hólar: [n.n.]. On the Bækur.is website: <http://baekur.is/> (June 27, 2012).
- . 1857. *Heilige und erbauliche Passions-Andachten. Aufs Neue herausgegeben*

- und durchgesehen*. Leipzig: Ernst Brecht, 1857. On the Google Books website: <http://books.google.co.uk/> (June 27, 2012).
- Lenneberg, Eric H. and John M. Roberts. 1956. *The Language of Experience: A Study in Methodology*. Baltimore: Waverly Press.
- Lindsey, Delwin T. and Angela M. Brown. 2006. Universality of Color Names. *Proceedings of the National Academy of Sciences of the United States of America* 103.44: 16608–16613.
- Magnús Stephensen. 1797. *Skemmtileg vinagleði*. [Skémtileg Vina-Gledi í fróðlegum Samræðum og Liódmælum]. Leirárgarðar: Björn Gottskálksson. On the Tímarit.is website: <http://timarit.is/> (June 27, 2012).
- OED = *The Oxford English Dictionary*. 1989. 20 vols. (2nd edition). Editors. J.A. Simpson and E.S.C. Weiner. Oxford: Clarendon Press; New York: Oxford University Press.
- OED online = *Oxford English Dictionary Online*. 2011. New York: Oxford University Press. On the website: <http://www.oed.com/> (June 27, 2012).
- ONP = *Ordbog over det norrøne prosasprog. A Dictionary of Old Norse Prose*. 2010. On the website: <http://www.onp.hum.ku.dk/> (June 27, 2012).
- Ordbog over det Danske Sprog*. 1975. 27 vols. (3rd edition) Copenhagen: Gyldendal.
- Páll Sveinsson. (ed.). 1862. Um framfarir heimsins frá því er Island fannst. *Ný sumargjöf* 4: 64–101. On the Tímarit.is website: <http://timarit.is/> (June 27, 2012).
- Poisson, Étienne Ljóni. *Um litarorðaförða í íslensku*. B.A. thesis. Hugvísindasvið, Háskóli Íslands. May 2011. On the Skemman.is website: <http://hdl.handle.net/1946/8415> (June 27, 2012).
- Regier, Terry, Paul Kay and Naveen Khetarpal. 2007. Color Naming Reflects Optimal Partitions of Color Space. *Proceedings of the National Academy of Sciences of the United States of America* 104.4: 1436–1441. On the JSTOR.org website: <http://www.jstor.org/stable/25426301> (June 11, 2012).
- ROH = *Ritmálssafn Orðabókar Háskólans*. On the Institute of Lexicography website: <http://www.lexis.hi.is/> (June 27, 2012).
- Root, Waverley. 1980. *Food: An Authoritative and Visual History and Dictionary of the Foods of the World*. New York: Simon and Schuster.
- Steingrímur Thorsteinsson. 1886. *Þýzk lestrarbók með stuttri málmyndalýsingu og orðasafni*. Reykjavík: [n.n.].
- Sverrir Hólmarsson, Christopher Sanders and John Tucker. 1989. *Íslensk-ensk orðabók/Concise Icelandic-English Dictionary*. On the website *Icelandic Online Dictionary and Readings*, University of Wisconsin Digital Collections: <http://digital.library.wisc.edu/1711.dl/IcelOnline> (June 27, 2012).
- Toussaint-Samat, Maguelonne. 1992. *A History of Food*. Trans. Anthea Bell. Cambridge: Blackwell Publishers.

- Vulgate = *Biblia sacra juxta vulgatam Clementiam. Editio Electronica*. 2005. Editor Michael Tweedale. On the VulSearch & The Clementine text project website: <http://vulsearch.sourceforge.net/html/index.html> (June 27, 2012).
- Wikipedia. *Frjálsa alfræðiritið*. "Appelsína." 2011. On the Wikipedia website: <http://is.wikipedia.org/wiki/Appelsína> (June 27, 2012).
- Witkowski, Stanley R. and Cecil H. Brown. 1977. An Explanation of Color Nomenclature Universals. *American Anthropologist* 79.1: 50–57.
- Wolf, Kirsten. 2005. Reflections on the Color of Esau's Pottage of Lentils (*Stjórn* 160.26–161.9). *Gripla* 16: 251–257.
- . 2006. Some Comments on Old Norse-Icelandic Color Terms. *Arkiv för nordisk filologi* 121: 173–92.
- . 2009. The Color Grey in Old Norse-Icelandic Literature. *Journal of English and Germanic Philology* 108.2: 222–238.
- . 2010. Towards a Diachronic Analysis of Old Norse-Icelandic Color Terms: The Cases of Green and Yellow. *Orð og Tunga* 12: 109–130.
- Þjóðviljinn* 282. December 31, 1951. On the Tímarit.is website: <http://timarit.is/> (June 27, 2012).
- Póra Andrea Nikólína Jónsdóttir. 1858. *Ný matreiðslubók ásamt ávísun um litun, þvott o.fl.*. Akureyri: Prentsmiðja Norður- og Austurumdæmisins. On the Bækur.is website: <http://baekur.is/> (June 27, 2012).

Keywords

Basic color terms, orange (fruit), orange (color), *appelsínugulur*.

Lykilorð

Grundvallarlitaorð, appelsína (ávöxtur), appelsínugulur (litur).

Útdráttur

Brent Berlin og Paul Kay halda því fram í *Basic Color Terms* (1969) að grundvallarorð yfir liti (*svartur, hvítur, rauður, grænn, gulur, blár, brúnn, fjólublár, appelsínugulur, grár*) séu tekin inn í tungumál eftir ákveðinni reglu í sjö þrepum. Kirsten Wolf (2006; 2010) hefur haldið því fram að nútímaíslenska sé á þrepi VII og hafi einungis níu grundvallarorð yfir liti og vanti því grundvallarorð fyrir 'purple' (*fjólublár*) og 'orange' (*appelsínugulur*). Þessi grein gefur yfirlit yfir þróun hugtakanna sem notuð hafa verið í íslensku fyrir appelsínuávöxtinn og appelsínugula litinn. Rannsakadár eru ástæður þess að Íslendingar tóku ekki grundvallarorðið 'orange' upp í mál sitt eins og önnur norræn tungumál, en völdu í staðinn að nota samsett orð tengt gula litnum (*appelsínugulur*). Höfundurinn heldur því ennfremur fram að þó að *appelsínugulur* hafi einu sinni verið skynjaður sem afbrigði af gulu hafi orðið náð sálfræðilega sjálfstæðri stöðu í málinu ('psychologically salient' á ensku) og ætti því að vera talið sem grundvallarlitarorð þó svo að það sé samsett. Því ætti nútímaíslenska að teljast vera á

Þrepi VII með að minnsta kosti tíu grundvallarlitarorð: *svartur, hvítur, rauður, grænn, gulur, blár, brúnn, grár, bleikur og appelsínugulur.*

*Susanne M. Arthur
Department of Scandinavian Studies
University of Wisconsin-Madison
1306 Van Hise Hall, 1220 Linden Drive
Madison, Wisconsin 53706, U.S.A.
sarthur@wisc.edu
s.arthur@twoclouds.net*



Kirsten Wolf

Basic Color Terms in Old Norse-Icelandic: A Quantitative Study

1 Introduction

In 1969, Brent Berlin and Paul Kay published their analysis of color terms in close to one hundred of the world's languages, arguing that there was a universal inventory of eleven basic color terms, located in the color space where English speakers place the most typical examples of black, white, red, orange, yellow, brown, green, blue, purple, pink, and grey. Comparing the vocabularies of languages possessing fewer than these eleven color terms, they maintained that basic color terms do not appear haphazardly in the diachronic development of a language, but that a language adds basic color terms in a particular order, which is interpreted as an evolutionary sequence:

- Stage I: black, white
- Stage II: black, white, red
- Stage IIIa: black, white, red, green
- Stage IIIb: black, white, red, yellow
- Stage IV: black, white, red, yellow, green
- Stage V: black, white, red, yellow, green, blue
- Stage VI: black, white, red, yellow, green, blue, brown
- Stage VII: black, white, red, green, yellow, blue, brown, purple, pink, orange, grey.

Since the publication of Berlin and Kay's study, there has been some elaboration and criticism of their conclusions, and in response to these, the two linguist-anthropologists have modified them. In 1975, *Orð og tunga* 15 (2013), 141–161. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

Brent Berlin and Elois Ann Berlin introduced a light-warm versus dark-cool stage instead of the earlier categorization based on brightness contrast. They also suggested that the dark-cool category would at stage IIIa divide, leading to the emergence of grue (blue or green), while the warm-light category would divide at stage IIIb, leading to the emergence of yellow. At stage IV, stage IIIa systems would then encode focal yellow, while stage IIIb would encode focal grue, which would find a focus in either blue or green. At stage V, undifferentiated grue would divide, and the new category would encode either blue or green. At stage VI, focal brown would emerge from the light-warm category. At stage VII, focal purple, pink, orange, and grey would become encoded, grey possibly early. In 1978, Paul Kay and Chad K. McDaniel argued that grey could occur at any stage from IIIa onward or potentially even earlier. In 1991, Paul Kay, Brent Berlin, and William Merrifield maintained that either brown or purple or both often appear before the green/blue composite is dissolved, and that there seems to be no fixed order in the temporal appearance of the two color terms. And in 2007, Terry Regier, Paul Kay, and Naveen Khetarpal argued against the view that color categories are organized around universal focal colors and in favor of K. A. Jamesson and R. G. D'Andrade's earlier proposal that "color naming is determined in part by general principles of categorization partitioning an irregularly shaped color solid" (1441).

Old Norse-Icelandic has eight basic color terms (*svartr*, *hvítr*, *rauðr*, *grænn*, *gulr*, *blár*, *brúnn*, *grár*), making it an early stage VII language (Wolf 2006). Recent analyses have largely supported Berlin and Kay's evolutionary sequence, though it has been argued that grey (*grár*) should be assigned an early stage, either stage III or stage IV (Wolf 2009), that blue (*blár*) should be assigned a late stage (Wolf 2006), that green (*grænn*) should be assigned a stage earlier than yellow (*gulr*) (Wolf 2009), and that brown (*brúnn*) should be assigned a fairly late stage, probably as late as blue (Wolf [forthcoming]).

The purpose of this article is a quantitative analysis of basic color terms in Old Norse-Icelandic literature in order to explore, if there is a correlation between the amended version for Old Norse-Icelandic of Berlin and Kay's sequence and the frequencies of *svartr*, *hvítr*, *rauðr*, *grænn*, *gulr*, *blár*, *brúnn*, and *grár*. It is, of course, recognized that there may not necessarily be a correlation between synchronic frequencies of usage and diachronic stage of introduction, and Berlin and Kay's original hypothesis does not predict such a correlation. After all, their

explanation is neuropsychological and not functional. What Berlin and Kay predict is that certain color differentiations are more readily perceived, but not that the environment of use necessarily provides more regular occasions for using the more salient terms. It should also be noted that near-synonyms, such as *døkk* and *myrkr* (for *svartr* and *blár*) and *ljóss* and *fagr* (for *hvítr*) have not been included, nor have the secondary color terms *blakkr*, *hárr*, *høss* (which occurs only in poetry), *jarpr*, *rjóðr*, and *bleikr*. The first five terms are collocationally restricted (Wolf 2006: 179–180), and *bleikr* should probably be regarded as a macrocolor covering, at least partly, the category of pale or light colors (Wolf 2005: 253 and 2006: 187).

The data are drawn from the following texts, which have all been excerpted (cf. *Íslenzk fornrit* 1933>): *Auðunar þáttur vestfirzka* (VI), *Bandamanna saga* (VII), *Bárðar saga Snæfellsáss* (XIII), *Bergbúa þáttur* (XIII), *Bjarnar saga Hítðelakappa* (III), *Brandkrossa þáttur* (XI), *Brands þáttur ǫrva* (IV), *Draumr Þorsteins Síðu-Hallssonar* (XI), *Droplaugarsona saga* (III), *Egils saga Skalla-Grimssonar* (II), *Egils þáttur Síðu-Hallssonar* (XIII), *Eiríks saga rauða* (IV), *Eyrbyggja saga* (IV), *Finnboga saga ramma* (XIV), *Fljótsdæla saga* (XI), *Flóamanna saga* (XIII), *Fóstbræðra saga* (VI), *Gísla saga Súrssonar* (VI), *Gísls þáttur Illugasonar* (III), *Grettis saga* (VII), *Grœnlendinga saga* (IV), *Grœnlendinga þáttur* (IV), *Gull-Ásu-Þórðar þáttur* (XI), *Gunnars saga Keldugnúpsfífls* (XIV), *Gunnars þáttur Þiðrandabana* (XI), *Gunnlaugs saga orms-tungu* (III), *Halldórs þættir Snorrasonar* (V), *Hallfreðar saga* (VIII), *Harðar saga ok Hólmverja* (XIII), *Hávarðar saga Ísfirðings* (VI), *Heiðarvígja saga* (III), *Hrafnkels saga Freysgoða* (XI), *Hrafnis þáttur Guðrúnarsonar* (VIII), *Hreiðars þáttur* (X), *Hrómundar þáttur halta* (VIII), *Hænsa-Þóris saga* (III), *Jökuls þáttur Búasonar* (XIV), *Kjalnesinga saga* (XIV), *Kormáks saga* (VIII), *Króka-Refs saga* (XIV), *Kumlbúa þáttur* (XIII) *Laxdæla saga* (V), *Ljósvetninga saga* (X), *Njáls saga* (XII), *Odds þáttur Ófeigssonar* (VII), *Ófeigs þáttur* (X), *Orms þáttur Stórolfssonar* (XIII), *Reykðæla saga ok Víga-Skútu* (X), *Sneglu-Halla þáttur* (IX), *Stjörnu-Odda draumr* (XIII), *Stúfs þáttur* (V), *Svarfdæla saga* (IX), *Sǫrla þáttur* (X), *Þórarins þáttur Neffjölfssonar* (XIII), *Þórðar saga hreðu* (XIV), *Þorgríms þáttur Hallasonar* (IX), *Þorleifs þáttur jarlsskálds* (IX), *Þormóðar þáttur* (VI), *Þorskfirðinga saga* (XIII), *Þorsteins saga hvíta* (IX), *Þorsteins saga Síðu-Hallssonar* (XI), *Þorsteins þáttur Austfirðings* (XI), *Þorsteins þáttur forvitna* (XIII), *Þorsteins þáttur stangarhöggs* (XI), *Þorsteins þáttur sogufróða* (XI), *Þorsteins þáttur tjaldstæðings* (XIII), *Þorsteins þáttur uxafóts* (XIII), *Þorvalds þáttur tasalda* (IX), *Þorvarðar þáttur krákunefs* (VI), *Valla-Ljóts saga* (IX), *Vápnfirðinga saga* (XI), *Vatnsdæla saga* (VIII), *Víga-Glúms saga* (IX), *Víglundar saga* (XIV), *Ögmundar þáttur dytts* (IX), and *Ölkofra þáttur* (XI).

Section 2 examines the basic color terms in the order of their frequency in the Sagas and *þættir* of Icelanders. Section 3 summarizes the data in illustrations and tables. Considerations and conclusions are presented in section 4.

2 Quantitative study

In the following, the eight Old Norse-Icelandic basic color terms are examined in the order of their frequency in the Sagas and *þættir* of Icelanders,¹ which is as follows: *rauðr*, *hvítr*, *blár*, *svartr*, *grár*, *grænn*, *brúnn*, and *gulr*. Bynames (which are marked with an asterisk in Tables 1–8) and toponyms are counted only once (with more than one occurrence marked with an “f” following the reference), and reference is made to the earliest occurrence of the name in a given saga or *þátr*. Personal and animal names are mentioned and referred to but excluded from the count, since naming traditions rather than color may have been a consideration, at least as far as personal names are concerned. Compounds such as *rauðbrúnn* or *gulgrænn* are listed under both *rauðr* and *brúnn* and *gulr* and *grænn*, respectively. Derived words, such as *hvítingr* and *blána*, are included, as are etymologically related words, such as *sorti*.

2.1 Rauðr

Rauðr is by far the most frequently used color term. Altogether, there are 184 occurrences. This is excluding the personal names *Rauðr*, *Rauðfeldr*, *Rauða-björn*, and also *Rauðgrani* (which appears as a name for Odin in *Bárðar saga Snæfellsáss* 163.5f) and *Heiðarauðr* (which appears as the name of a horse in *Fljótsdæla saga* 256.17). Of these 184 occurrences, 48 are compounds, 13 being toponyms.

1 References in the following are all to the *Íslenzk fornrit* editions. Some of the texts contain no color terms. These comprise *Bergbúa þátr*, *Brands þátr ǫrva*, *Draumr Þorsteins Síðu-Hallssonar*, *Egils þátr Síðu-Hallssonar*, *Gull-Ásu-Þórðar þátr*, *Stjörnu-Odda draumr*, *Stúfs þátr*, *Þormóðar þátr*, *Þorsteins þátr Austfirðings*, *Þorsteins þátr forvitna*, *Þorsteins þátr sögufróða*, and *Qgmundar þátr dytts*. References are to page and line numbers. In cases where two versions of a given text are presented synoptically, reference is given to both, e.g., *hvítaváðir* in *Víga-Glúms saga* 97.33/98.1.

Literary work	N	References	Compounds	Toponyms
<i>Bárðar saga Snæfellsáss</i>	2	117.18, 126.9	<i>rauðskeggjaðr</i> (126.9)	Rauðfeldsgjá (117.18)
<i>Bjarnar saga Hitdælakappa</i>	4	148.17, 187.1, 193.12, 197.12	<i>barðrauðr</i> (193.12), <i>rauðskeggjaðr</i> (197.12)	
<i>Brandkrossa þátr</i>	1	189.1		
<i>Droplaugarsona saga</i>	1	*141.2		
<i>Egils saga Skalla-Grímssonar</i>	9	41.24, *57.1, 64.16, 78.15, 98.9, *129.4f, 170.21 (<i>rauða</i>), 188.13, 243.30	<i>dreyrrauðr</i> (64.16), <i>rauðablástr</i> (78.15)	Rauðsgil (98.9)
<i>Eiríks saga rauða</i>	5	195.10f, 224.19, 228.9, 13.25	<i>rauðskeggjaðr</i> (224.19)	
<i>Eyrbyggja saga</i>	8	*12.23f, 20.16f, 21.5, 26.22, 47.9, *59.11f, 99.11f, 120.21	<i>rauðbleikr</i> (21.5), <i>rauðskeggjaðr</i> (26.22)	Rauðamelr (20.16f), Rauðavíkrhøfði (99.16f)
<i>Finnboga saga ramma</i>	2	276.10, 282.10		
<i>Fljótsdæla saga</i>	4	225.15, 227.10, 237.3-4, 256.17	<i>rauðlitaðr</i> (237.3-4)	
<i>Flóamanna saga</i>	7	239.1, *245.10, 270.12/27, 278.5/18, 292.13, 300.16/31, *301.7	<i>rauðskeggjaðr</i> (278.5/18)	Rauða (239.1f)
<i>Fóstbræðra saga</i>	7	153.15, 154.8, 266.20, 271.9, 275.3, 276.11, 15		
<i>Gísla saga Súrssonar</i>	3	*16.6, 103.8, 12	* <i>rauðfeldr</i> (16.6), <i>blóðrauðr</i> (103.12)	
<i>Gísls þátr Illugasonar</i>	1	334.14		
<i>Grettis saga</i>	8	27.27, 36.22, 86.15, *91.9, 96.29, *180.9f, 250.1, 265.24	<i>rauðhærðr</i> (36.22)	Rauðahaf (86.15)
<i>Gunnars saga Keldugnúpsfjfls</i>	2	350.19, 370.22		
<i>Gunnlaugs saga ormstungu</i>	4	61.17f, 65.8, 70.1, 74.2		Rauðamelr (61.17f)
<i>Hallfreðar saga</i>	2	171.22-172.1/172.6, 175.15 (<i>rauða</i>)	<i>rauðskeggjaðr</i> (171.22-172.1)	
<i>Harðar saga ok Hólmverja</i>	1	59.6	<i>tandrauðr</i>	
<i>Hávarðar saga Ísfirðings</i>	1	336.12f		Rauðasandr
<i>Heiðarvíga saga</i>	5	237.8, 247.4, 278.4, 319.16, *326.19	<i>alrauðr</i> (237.8)	
<i>Hrafnkels saga Freysgoða</i>	1	111.4	<i>rauðlitaðr</i>	
<i>Hrómundar þátr halta</i>	2	310.15, 313.8		

Literary work	N	References	Compounds	Toponyms
<i>Hænsa-Þóris saga</i>	1	23.4		
<i>Kjalnesinga saga</i>	2	29.29, 43.15		
<i>Kormáks saga</i>	5	214.8, 217.7, 249.16, 261.23 (rauða), 286.10	<i>mórauðr</i> (214.8), <i>hvarmrauðr</i> (217.7), <i>mosrauðr</i> (249.16)	
<i>Króka-Refs saga</i>	1	134.22		
<i>Kumlbúa þátr</i>	1	454.1		
<i>Laxdæla saga</i>	17	21.3, 55.16, 57.29, 79.14, 91.2-3, 118.7, 127.27, 134.31, 135.13, 139.19, 168.8, 187.12, 189.4, 20, 194.16, 225.11, 13	<i>*rauðnefr</i> (21.3), <i>deyrrauðr</i> (91.2-3), <i>alrauðr</i> (127.27), <i>rauðleitr</i> (189.20), <i>rauðbrúnn</i> (194.16)	
<i>Ljósvetninga saga</i>	3	14.9, 79.1, 85.21		
<i>Njáls saga</i>	25	39.20, 44.3, 64.20, 85.20, 92.13f, 114.15, 115.22, *146.7, 147.16, 150.17, 187.2, 228.27, 231.20x2, *292.2, 8, *309.22, *347.4f, 351.25, 352.1, 361.6, 409.15, 448.21f, 454.17, 467.11	<i>deyrrauðr</i> (39.20), <i>rauðálfr</i> (115.22)	Rauðaskriða (92.13f)
<i>Orms þátr Ófeigssonar</i>	2	121.19, 373.21		
<i>Reykðæla saga ok Víga-Skútu</i>	4	153.1f, 177.5, 221.17, 222.4	<i>alrauðr</i> (222.4)	Rauðaskriða (153.1f)
<i>Sneglu-Halla þátr</i>	3	263.8/264.26, 271.27, 275.5		
<i>Svarfdæla saga</i>	6	136.15, 151.1-2f, 153.10f, 169.14, 174.13, 184.20	<i>*rauðfeldr</i> (153.10f)	
<i>Sþrla þátr</i>	1	113.6		Rauðalækur
<i>Þórarins þátr Nefjölfssonar</i>	1	336.21		
<i>Þórðar saga hreðu</i>	4	168.9, 176.25, 192.4, 195.16		
<i>Þorleifs þátr jarlsskálds</i>	1	219.4		
<i>Þorskfirðinga saga</i>	5	*175.6-7, *177.3, 8f, 184.12, 205.9		Rauðsdalr (177.8f)
<i>Þorsteins þátr stangarhöggs</i>	1	69.3	<i>rauðvíkingr</i>	
<i>Þorsteins þátr tjaldstæðings</i>	1	*430.9		
<i>Þorsteins þátr uxafóts</i>	2	352.9, 22	<i>rauðklæddr</i>	
<i>Þorvarðar þátr krákunefs</i>	1	372.1		

Literary work	N	References	Compounds	Toponyms
<i>Valla-Ljóts saga</i>	1	243.9		
<i>Vápnfirðinga saga</i>	8	*23.11, 48.22, 24, 25, 28, 49.2-3, 5, 51.17	<i>rauðfleckótr</i> (48.22, 25, 49.2-3)	
<i>Vatnsdæla saga</i>	4	68.21, 111.3, 4, 7		
<i>Víga-Glúms saga</i>	3	37.9, 51.11, 86.7	* <i>rauðkinnr</i> (37.9)	Rauða-hjalli (51.11)
<i>Víglundar saga</i>	1	108.5		

Table 1. Occurrences of *rauðr*

2.2 Hvíttr

Hvíttr is the second most commonly used color term. There are 150 occurrences. Of these, 33 are bynames, two being compound bynames. The adjective appears commonly in toponyms.

Literary work	N	References	Compounds	Toponyms
<i>Bandamanna saga</i>	1	297.11		Hvítá
<i>Bárðar saga Snæfellsáss</i>	3	102.16, 17, 170.9	<i>hvítaváðir</i> (170.9)	
<i>Bjarnar saga Híttdælakappa</i>	10	122.17, 136.10x2, 140.4, 144.13, 187.1, 2, 189.19, 195.7f, 198.2	<i>alhvíttr</i> (136.10, 187.1), <i>hvítmael</i> (189.19)	Hvítá (122.17), Hvít-ingshjalli (195.7f)
<i>Droplaugarsona saga</i>	1	*143.12f		
<i>Egils saga Skalla-Grímssonar</i>	7	*16.9f, 74.9f, 99.8, 219.30, 274.3, *291.19, 299.9 (<i>hvítna</i>)	* <i>hvítbeinn</i> (219.30)	Hvítá (74.9f), Hvítárvellir (99.8)
<i>Eiríks saga rauða</i>	6	*195.2, 206.21, 207.3, 227.13, 234.3, 234.5		Hvíttra-mannaland (234.5)
<i>Eyrbyggja saga</i>	8	*4.22f, 74.4 (<i>hvítigr</i>), 108.18, 120.24, 127.16, *133.7f, 143.25f, 178.2	<i>fannhvíttr</i> (108.18)	Hvítá (143.25f)
<i>Finnboga saga ramma</i>	1	*286.18f		
<i>Fljótsdæla saga</i>	3	237.2, 261.13, 274.9-10		
<i>Flóamanna saga</i>	2	*242.7, *258.5f		
<i>Fóstbræðra saga</i>	15	*124.12, 128.25f, 176.6 (<i>hvítigr</i>), 178.3f, *181.16, 181.18, 195.1, 231.5, 233.16, 234.1, 12, 270.7 (<i>hvítigr</i>), 276.11, 13, 17		Hvítá (128.25f), Hvítstaðir (178.3f)

Literary work	N	References	Compounds	Toponyms
<i>Gísla saga Súrssonar</i>	2	48.2, 117.13f		Hvítá (117.13f)
<i>Grettis saga</i>	5	27.24, 48.19-20, 21f, 54.18, *209.18f	<i>fránhvítigr</i> (27.24)	Hvítársíða (48.19-20), Hvítá (48.21f)
<i>Grænendinga saga</i>	3	250.3, 255.4, 276.12		Hvítserkr (276.12)
<i>Grænendinga þátr</i>	1	291.1	<i>hvítabjörn</i>	
<i>Gunlaugs saga ormstungu</i>	6	51.12, 57.22, 24, 58.13, 68.16, 90.5		Hvítársíða (58.13)
<i>Hallfredar saga</i>	5	150.14, 170.17, 171.12, 182.12, 198.1	*Hvíta-Kristr (170.17)	Hvítá (150.14)
<i>Harðar saga ok Hólmverja</i>	5	26.4, 32.22, 50.10f, 77.12, 96.15	<i>hvítabjörn</i> (77.12)	Hvítársíða (26.4), Hvítá (50.10f), Hvítárvellir (96.15)
<i>Hávarðar saga Ísfirðings</i>	1	343.8		
<i>Heiðarvíga saga</i>	6	230.18f, 247.14f, 260.9, 272.16, 291.12, 294.18	<i>hvítvöndr</i> (291.12)	Hvítársíða
<i>Hænsa-Þóris saga</i>	1	36.16		Hvítá
<i>Kjalnesinga saga</i>	1	31.17		
<i>Kormáks saga</i>	2	211.4, 265.1	<i>alhvítr</i> (211.4)	
<i>Króka-Refs saga</i>	8	133.27, 134.10, 142.7, 143.1, 151.13, 157.7, 8, 158.7	<i>hvítabjörn</i> (133.27, 134.10, 142.7, 143.1, 157.7)	
<i>Laxdæla saga</i>	10	*66.21f, 123.16, 131.12, 132.16f, 135.13, 148.5f, 158.16, 160.8, 188.11, 233.21	<i>hvítaváðir</i> (123.16, 158.16)	Hvítá (132.16f), Hvítadalr (148.5f)
<i>Njáls saga</i>	13	10.2f, *74.13-140, *120.13-14f, *181.17, *232.6f, *238.3, 246.16f, *256.6, *352.8, *424.10, *440.2,* 451.7, 461.17-18		Hvítá (10.2f), Hvítanes (246.16f), Hvítborg (461.17-18)
<i>Odds þátr Ófeigssonar</i>	1	373.21		
<i>Orms þátr Ófeigssonar</i>	1	411.4	<i>svanhvítr</i>	
<i>Reykðæla saga ok Víga-Skútu</i>	4	178.7, 181.16, 182.21, 233.3		
<i>Svarfðæla saga</i>	2	181.23, *211.25	<i>hvítabjörn</i> (181.23)	

Literary work	N	References	Compounds	Toponyms
<i>Þórðar saga hreðu</i>	4	*176.8, 178.28f, 180.15, 197.20		Hvítá (178.28f), Hvítárvellir (180.15)
<i>Þorskfirðinga saga</i>	3	*177.1, 197.20, 21		
<i>Þorsteins saga hvíta</i>	3	*3.2, *3.9-10f, *4.5f		
<i>Þorsteins þátr Síðu-Hallssonar</i>	1	314.15	<i>hvítabjörn</i>	
<i>Þorsteins þátr uxafóts</i>	2	*343.7, 350.14	<i>hvítabjarnarhúnn</i> (350.14)	
<i>Þorvarðar þátr tasalda</i>	1	126.28	<i>hvítaváðir</i>	
<i>Vatnsdæla saga</i>	4	7.11, 15, 50.3f, 126.7	<i>hvítaváðir</i> (126.7)	Hvítá (50.3f)
<i>Víga-Glúms saga</i>	4	15.19, 52.4/16, *57.10/29, 97.33/98.1	<i>hvítaváðir</i> (97.33/98.1)	
<i>Víglundar saga</i>	4	72.2, 99.9, 104.17, 110.12	<i>drifhvítr</i> (99.9), <i>mjallhvítr</i> (110.12)	

Table 2. Occurrences of *hvítr*

2.3 Blár

There are 138 occurrences of *blár*.² Almost half are compounds, five being bynames and 12 toponyms.

Literary work	N	References	Compounds	Toponyms
<i>Bandamanna saga</i>	1	329.8		Bláskógaheiðr
<i>Bárðar saga Snæfellsáss</i>	3	118.17, 131.6, 163.6	* <i>blátannarskegg</i> (118.17), * <i>bláfellingr</i> (131.6), <i>bláflekkóttir</i> (163.6)	
<i>Bjarnar saga Hítðalakappa</i>	4	136.19, 137.5, 154.18, 177.23		
<i>Egils saga Skalla-Grímssonar</i>	5	41.24, 142.19, 187.20, 210.11, 284.1		
<i>Eiríks saga rauða</i>	3	200.3f, 206.19, 229.2		Bláserkr (200.3f)
<i>Eyrbyggja saga</i>	7	23.3, 53.12, 78.10, 93.13, 134.3, 146.22, 169.30	<i>kolblár</i> (93.13, 146.22)	
<i>Finnboga saga ramma</i>	2	283.14, 20	<i>blámaðr</i>	
<i>Fljótsdæla saga</i>	1	286.31		

² For a discussion of *blár*, see Wolf (2006).

Literary work	N	References	Compounds	Toponyms
<i>Flóamanna saga</i>	3	246.22, 262.20, 275.5/20		
<i>Fóstbræðra saga</i>	2	128.1 (<i>blána</i>), 239.16		
<i>Gísla saga Súrssonar</i>	6	41.17, 52.26, 64.20, 65.22, 68.9, 82.8	<i>blákápuaðr</i> (65.22), <i>bláserkr</i> (68.9), <i>bláföld</i> (82.8)	
<i>Grettis saga</i>	5	112.9, 150.8, 213.4, 236.3, 252.7	<i>kolblár</i> (252.7)	
<i>Gunnars saga Keldugnúpsfífls</i>	9	266.16/25, 26, 367.5, 7, 14/22, 23-4, 29, 368.4, 6	<i>blámaðr</i>	
<i>Gunnlaugs saga ormstungu</i>	1	77.24		
<i>Halldórs þátr I</i>	1	257.2		
<i>Harðar saga ok Hólmverja</i>	7	32.23-4, 35.18, 39.22, 46.7f, 62.23f, 64.9, 65.16f	<i>bláeygr</i> (32.23-4), <i>blárendr</i> (39.22), <i>*blásiða</i> (46.7f), <i>*bláskeggr</i> (62.23f)	Bláskeggs-áróss (64.9), Bláskeggsá (65.16f)
<i>Hávarðar saga Ísfirðings</i>	2	292.2f, 327.11		Blámýri (292.2f)
<i>Heiðarvíga saga</i>	1	289.18		
<i>Hrafnkels saga Freysgoða</i>	3	104.23, 109.11, 128.5		Bláfjöll (109.11)
<i>Hrómundar þátr halta</i>	1	310.24	<i>bláfjallaðr</i>	
<i>Jökuls þátr Búasonar</i>	2	49.21, 59.10f		Bláland (59.10f)
<i>Kjalnesinga saga</i>	12	35.4, 26, 36.10, 24, 31, 37.2, 5, 8, 11, 12, 16, 42.25	<i>blámaðr</i>	
<i>Kormáks saga</i>	3	246.5, 260.13, 269.19	<i>bláfiðraðr</i> (260.13)	Bláland (269.19)
<i>Króka-Refs saga</i>	4	130.1 (<i>blána</i>), 151.11-2, 152.2, 154.8		
<i>Laxdæla saga</i>	12	82.11, 95.14, 168.14, 185.2, 187.7, 9, 16, 188.23, 198.14, 224.12, 237.3, 245.7	<i>kolblár</i> (82.11), <i>bláeygr</i> (187.16)	Bláskógaheiðr (95.14)
<i>Ljósvetninga saga</i>	1	88.11		
<i>Njáls saga</i>	16	35.16, 44.3, 53.10, 63.6f, 78.6, 146.4-5, 227.10, 231.15, 292.3, 296.12, 304.2x2, 314.11, 335.1 (<i>blánaðr</i>), 348.7, 421.14	<i>kolblár</i> (35.16, 78.6), <i>bláeygr</i> (53.10), <i>*blátannarskegg</i> (146.4-5), <i>blárendr</i> (304.2)	
<i>Ófeigs þátr</i>	1	119.13		
<i>Svarfáela saga</i>	1	136.16		
<i>Þórðar saga hreðu</i>	1	176.25		
<i>Þorskfirðinga saga</i>	2	183.11, 220.26		
<i>Þorsteins þátr uxafóts</i>	3	352.24, 354.2, 360.18	<i>bláklæddr</i> (352.24, 354.2)	

Literary work	N	References	Compounds	Toponyms
<i>Borvalds þátr tasalda</i>	1	124.29	<i>blámaðr</i>	
<i>Valla-Ljóts saga</i>	3	235.29, 240.18, 245.1		
<i>Vatnsdæla saga</i>	2	60.26, 101.6–7		Bláskógaheiðr (101.6–7)
<i>Víga-Glúms saga</i>	4	16.21, 27.3/18, 86.14, 92.27	<i>alblár</i> (92.27)	
<i>Víglundar saga</i>	2	78.8, 17	<i>bláklæddr</i>	
<i>Ólkofra þátr</i>	1	83.1–2f		Bláskógar

Table 3. Occurrences of *blár*

2.4 Svarttr

There are 109 occurrences of *svarttr*. This is excluding the personal names *Svarttr*, *Svarthöfði*, *Svertingr*, *Surtr*, and also *Svartfaxi*, which appears as the name of a horse in *Harðar saga ok Hólmverja* (10.16f). Of these, 17 are compounds, one being a toponym.³

Literary work	N	References	Compounds	Toponyms
<i>Bandamanna saga</i>	4	318.12/25, 315.17, 354.24/355.3, 363.16	<i>svartleggja</i> (354.24/355.3)	
<i>Bárðar saga Snæfellsáss</i>	5	105.24, *111.1, 148.6, 160.10, 162.21	<i>svartskeggjaðr</i> (160.10)	
<i>Bjarnar saga Hítðælakappa</i>	3	187.3, 200.18, 204.11		
<i>Droplaugarsona saga</i>	2	162.1, *163.15	<i>svartaskáld</i> (163.15)	
<i>Egils saga Skalla-Grímssonar</i>	7	5.7, *7.2f, 80.20, 96.24, 143.19, *151.19, 194.14	<i>svarteygr</i> (143.19), <i>svartbrúnn</i> (194.14)	
<i>Eiríks saga rauða</i>	3	206.20, 222.2, 227.16		
<i>Eyrbyggja saga</i>	6	*14.17f, 23.5x2, 139.9, 140.7, 184.5	<i>svartbrúnn</i> (139.9)	
<i>Finnboga saga ramma</i>	1	327.4		
<i>Fljótsdæla saga</i>	5	238.1, 248.18, 249.32, 279.9, 294.16 (<i>sorti</i>)		
<i>Flóamanna saga</i>	1	291.5/19	<i>svartbaksegg</i>	
<i>Fóstbræðra saga</i>	12	124.2, *150.9f, 160.9, 170.6, 231.5, 28, 235.4, 16, 236.1, 237.2, 250.20, 254.4 (<i>sorti</i>)		

³ In the example from *Svarfdæla saga*, the color reference in the phrase “[þ]ar töldu þeir fimmtán skip hjá drekanum, ok var tjaldat yfir svörtu” could be an ellipsis for something like “svörtu klæði/tjaldi.”

Literary work	N	References	Compounds	Toponyms
<i>Gísla saga Súrssonar</i>	1	15.18		
<i>Grettis saga</i>	5	148.11, 149.16, 150.11-2, 151.6, 217.11		
<i>Grænendinga saga</i>	2	*258.5f, 262.23		
<i>Gunnars saga Keldugnúpsfífls</i>	2	358.20, 360.25		
<i>Gunnars þátr Þiðrandabana</i>	1	196.6		
<i>Gunnlaugs saga ormstungu</i>	5	54.11, 59.3, *62.11f, 69.17, 96.12	<i>svarteygr</i> (54.11, 59.3)	
<i>Harðar saga ok Hólmverja</i>	2	46.4, 54.22		
<i>Heiðarvíga saga</i>	2	260.10, 272.17		
<i>Hænsa-Þóris saga</i>	1	29.9		
<i>Kjalnesinga saga</i>	1	9.19		
<i>Kormáks saga</i>	5	206.2, 210.11, 211.9, 11, 288.20		
<i>Laxdæla saga</i>	8	*66.22f, 101.8, *164.3f, 188.10, 18, 189.14, 20, 245.6	<i>svarteygr</i> (189.14), <i>svartjarpr</i> (189.20)	
<i>Njáls saga</i>	10	*36.13, 38.8, 95.13, 133.11, *181.27f, *271.1, 304.3, 321.2, *437.18, 446.14	<i>alsvartr</i> (133.11)	
<i>Ófeigs þátr</i>	1	121.21	<i>alsvartr</i>	
<i>Orms þátr Stórolfssonar</i>	3	407.15, 409.17, 415.16	<i>kolsvartr</i>	
<i>Reykðæla saga ok Víga-Skútu</i>	1	233.3		
<i>Svarfæla saga</i>	2	136.3 (<i>svarta</i>), 162.16		
<i>Þorsteins þátr uxafóts</i>	1	360.18		
<i>Valla-Ljóts saga</i>	2	247.11, 256.4	<i>svartklæddr</i> (247.11)	
<i>Vatnsdæla saga</i>	3	73.4, 106.4, 120.22		Svartsfells- mýrar (106.4)
<i>Víga-Glúms saga</i>	1	52.4/16		
<i>Víglundar saga</i>	1	84.10 (<i>sorti</i>)		

Table 4. Occurrences of *svarttr*

2.5 Grár

Grár is less common than *svarttr*. There are 69 occurrences of the color term in the texts examined. Almost half of these are compounds, two being toponyms. Examples of *grár* (cf. *gráleikr*, *gráligr*, *grálíga*, *grályndr*)

in the meaning “hostile” have not been included, nor have examples of *grár* as a byname, since the adjective may describe an individual’s unpleasant personal traits rather than aspects of his physical appearance.⁴

Literary work	N	References	Compounds	Toponyms
<i>Bárðar saga Snæfellsáss</i>	6	127.16, 129.9, 133.3, 135.2, 139.6, 145.21		
<i>Bjarnar saga Híttdælakappa</i>	3	168.22, 169.4, 200.9f	<i>grámagi</i> (168.22, 169.4)	Grásteinn (200.9f)
<i>Droplaugarsona saga</i>	1	176.5		
<i>Egils saga Skalla-Grimssonar</i>	8	34.23, 42.2, 43.16, 70.10, 143.16, 219.14f, 233.4-5, 260.11	<i>grávara</i> (34.23, 42.2, 43.16, 233.4-5), <i>úlfrgrár</i> (143.16, 260.11), * <i>gráfeldr</i> (219.14f)	
<i>Eiríks saga rauða</i>	1	228.1	<i>algrár</i>	
<i>Eyrbyggja saga</i>	3	158.5, 170.23, 171.17	<i>apalgrár</i> (170.23, 171.17)	
<i>Fljótsdæla saga</i>	4	261.12, 256.16, 270.7, 286.31	<i>járngrár</i> (270.7)	
<i>Fóstbræðra saga</i>	1	147.1		
<i>Gísla saga Súrssonar</i>	5	5.5f, 90.9, 94.11, 21, 122.13	Grásíða (name of sword, 5.5f)	
<i>Gísls þátr Illugasonar</i>	1	334.14		
<i>Grettis saga</i>	3	110.3, 4, 187.2	<i>gráeygr</i> (110.3), <i>úlfrgrár</i> (110.4)	
<i>Grænendinga saga</i>	1	261.3	<i>grávara</i>	
<i>Gunnlaugs saga ormstungu</i>	3	63.6, 65.11, 68.16	<i>grásími</i> (63.6)	
<i>Hallfredar saga</i>	1	182.15		
<i>Harðar saga ok Hólmverja</i>	1	76.2	<i>apalgrár</i>	
<i>Hávarðar saga Ísfirðings</i>	1	309.2		
<i>Heiðarvíga saga</i>	4	217.7, 231.23, 232.2, 233.2	<i>gráfeldr</i> (217.7)	
<i>Hreiðars þátr</i>	1	250.29		
<i>Laxdæla saga</i>	6	79.14, 84.6, 104.3, 185.3, 188.19, 230.16	<i>gráfeldr</i> (79.14), <i>apalgrár</i> (84.6),	
<i>Njáls saga</i>	4	288.27, 320.23, 449.10, 454.14	<i>apalgrár</i> (449.10)	
<i>Orms þátr Stórolfssonar</i>	1	401.5		
<i>Reykðæla saga</i>	2	161.15, 221.25	<i>grábildótr</i> (161.15)	
<i>Svarfdæla saga</i>	2	132.4, 189.18	<i>gráskinn</i> (132.4)	
<i>Þórðar saga hreðu</i>	1	169.16	<i>gráskinn</i>	

4 For a discussion of *grár* and also *hárr*, a secondary color term, see Wolf (2009).

Literary work	N	References	Compounds	Toponyms
<i>Dorskfirdinga saga</i>	2	205.2, 233.19	<i>grákkollóttir</i>	Grásteinsdæld (233.19)
<i>Dorsteins þáttur uxafóts</i>	1	359.3	<i>grábjörn</i>	
<i>Dorvarðar þáttur krákunefs</i>	1	373.2	<i>algrár</i>	
<i>Víga-Glúms saga</i>	1	71.15		

Table 5. Occurrences of *grár*

2.6 Grænn

There are only 27 occurrences of *grænn*, of which 17 are compounds, 15 being toponyms.⁵

Literary work	N	References	Compounds	Toponyms
<i>Auðunar þáttur</i>	1	361.14f		Grænland
<i>Bárðar saga Snæfellsáss</i>	2	114.21f, 140.18f		Grænland (114.21f), Grænamýrartunga (140.18f)
<i>Eiríks saga rauða</i>	1	201.11f		Grænland
<i>Eyrbyggja saga</i>	1	60.11f		Grænland
<i>Finnboga saga ramma</i>	2	272.16f, 280.29		Grænmór (272.16f)
<i>Fljótsdæla saga</i>	1	227.2		
<i>Flóamanna saga</i>	2	236.14, 258.11f		Grænland (258.11f)
<i>Fóstbræðra saga</i>	2	204.17f, 276.12		Grænland (204.17f)
<i>Gísla saga Súrssonar</i>	1	118.2		Grænland
<i>Grænlandinga saga</i>	1	242.20f		Grænland
<i>Grænlandinga þáttur</i>	1	273.3f		Grænland
<i>Hallfreðar saga</i>	1	153.6		
<i>Hrafnkels saga Freysgoða</i>	1	111.3	<i>laufgrænn</i>	
<i>Hrafnar þáttur Guðrúnarsonar</i>	1	320.5		Grænteigr
<i>Jökuls þáttur Búasonar</i>	1	95.13		Grænland

5 For a discussion of *grænn*, see Wolf (2009).

Literary work	N	References	Compounds	Toponyms
<i>Kröka-Refs saga</i>	2	131.9f, 132.7		Grænland (131.9f)
<i>Laxdæla saga</i>	1	188.2	<i>gulgrænn</i>	
<i>Reykðæla saga ok Víga-Skútu</i>	1	233.1		
<i>Þóraríns þáttur Nefjólfssonar</i>	1	146.5		Grænland
<i>Víga-Glúms saga</i>	3	20.10, 52.2/14, 64.3		

Table 6. Occurrences of *grænn*

2.7 Brúnn

Brúnn is also uncommon.⁶ There are only 18 occurrences, and of these, six are compounds, two being toponyms.

Literary work	N	References	Compounds	Toponyms
<i>Bandamanna saga</i>	1	325.16		
<i>Eyrbyggja saga</i>	1	34.26	<i>möbrúnn</i>	
<i>Fimboga saga ramma</i>	1	256.5	<i>brúngras</i>	
<i>Fljótsdæla saga</i>	1	227.2		
<i>Flóamanna saga</i>	1	265.14		Brúnahaugr
<i>Gísls þáttur Illugasonar</i>	1	334.14		
<i>Grettis saga</i>	3	99.6, 148.7, 13		
<i>Hrafnkels saga Freysgoða</i>	1	100.15	<i>brúnmóalóttur</i>	
<i>Kormáks saga</i>	1	291.12	<i>brúnleggr</i>	
<i>Laxdæla saga</i>	1	194.16	<i>rauðbrúnn</i>	
<i>Njáls saga</i>	3	148.5, 276.22, 277.14		
<i>Þorgríms þáttur Hallasonar</i>	1	299.3f		Brúnastaðir
<i>Valla-Ljóts saga</i>	1	240.21-2		
<i>Víglundar saga</i>	1	77.3		

Table 7. Occurrences of *brúnn*

In addition, *Brúnn* occurs as the name of a horse (*Víglundar saga* 79.6), and *Brúni* appears as a personal name (*Heiðarvíga saga* 288.22, *Ljósveitninga saga* 16.15f, *Njáls saga* 407.20, and *Vápnfirðinga saga* 59.1).

⁶ For a discussion of *brúnn* and also *jarpr*, a secondary color term, see Wolf [forthcoming]. Examples of *svartbrúnn* and *skolbrúnn* have not been included, since *brúnn* may be referring to eyebrows and not color.

2.8 Gulr

Gulr is rare.⁷ The texts examined yield only five occurrences of the color term.

Literary work	N	References	Compounds	Toponyms
<i>Eyrbyggja saga</i>	1	78.9		
<i>Fóstbræðra saga</i>	1	276.11		
<i>Laxdæla saga</i>	2	187.14, 188.2	<i>gulgrænn</i> (188.2)	
<i>Njáls saga</i>	1	53.11		

Table 8. Occurrences of *gulr*

3 Summary of data

The relative frequency of the eight basic color terms in the Sagas and *þættir* of Icelanders may be illustrated as follows (Figure 1):

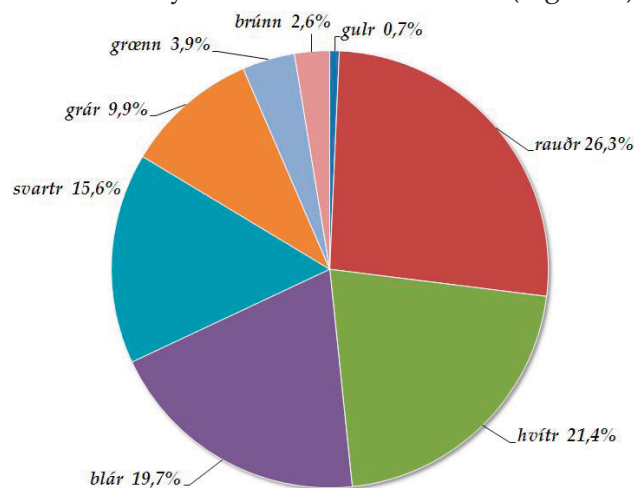


Figure 1. Relative frequency of basic color terms

If toponyms are excluded, *blár* is more common than *hvítr*, and *brúnn* is more common than *grænn*, and the sequence for the four color terms is then as follows: *blár* (20.1%) > *hvítr* (19.3%) > *brúnn* (2.6%) > *grænn* (1.9%). The sequence is the same, if both toponyms and bynames (including compound bynames) are excluded: *blár* (21.8%) > *hvítr* (17.4%) > *brúnn* (2.9%) > *grænn* (2.0%). If toponyms, bynames,

⁷ For a discussion of *gulr*, see Wolf (2009).

all compounds, and derivations of the color adjectives (that is, nouns and verbs) are excluded, the sequence also remains the same, except that *svartr* is more common than *blár* and *hvítr*. The sequence for all the eight basic color terms is then as follows: *rauðr* (28.9%) > *svartr* (18.6%) > *blár* (18.3%) > *hvítr* (18.0) > *grár* (9.8 %) > *brúnn* (2.8%), *grænn* (2.6%) > *gulr* (1.0%).

Njáls saga and *Laxdæla saga* have more than 50 occurrences of color terms and are followed by *Egils saga Skalla-Grímssonar* and *Fóstbræðra saga* with 36–40 occurrences. They are also among the longest sagas. *Bárðar saga Snæfellsáss*, *Bjarnar saga Hitdælakappa*, *Eiríks saga rauða*, *Eyrbyggja saga*, and *Grettis saga* have between 21 and 35 occurrences of color terms. *Fljótsdæla saga*, *Flóamanna saga*, *Gísla saga Súrssonar*, *Gunnars saga Keldugnúpsfífls*, *Gunnlaugs saga ormstungu*, *Harðar saga ok Hólmverja*, *Heiðarvíga saga*, *Kjalnesinga saga*, *Kormáks saga*, *Króka-Refs saga*, *Reykðæla saga ok Víga-Skútu*, *Svarfdæla saga*, *Þorskfirðinga saga*, *Vatnsdæla saga*, and *Víga-Glúms saga* have between 11 and 20 occurrences of color terms. Sagas with 6–10 occurrences comprise *Bandamanna saga*, *Finnboga saga ramma*, *Grænlandinga saga*, *Hallfreðar saga*, *Hrafnkels saga Freysgoða*, *Þórðar saga hreðu*, *Valla-Ljóts saga*, *Vápnfirðinga saga*, and *Víglundar saga*. *Droplaugarsona saga*, *Hávarðar saga Ísfirðings*, *Hænsa-Þóris saga*, *Ljósvefninga saga*, *Þorsteins saga hvíta*, and *Þorsteins saga Síðu-Hallssonar* have five or fewer occurrences; the four last-mentioned sagas are also the ones with the least variety of color terms.

Only *Eyrbyggja saga* and *Laxdæla saga* have all the eight color terms, and this is if toponyms, bynames, all compounds, and derivations of the color adjectives are included. In *Fljótsdæla saga*, *Fóstbræðra saga*, and *Njáls saga*, seven color terms are represented. In *Bárðar saga Snæfellsáss*, *Eiríks saga rauða*, *Finnboga saga ramma*, *Flóamanna saga*, *Gísla saga Súrssonar*, and *Víga-Glúms saga*, six color terms are represented. Sagas with five different color terms comprise *Bjarnar saga Hitdælakappa*, *Egils saga Skalla-Grímssonar*, *Grettis saga*, *Harðar saga ok Hólmverja*, *Heiðarvíga saga*, *Kormáks saga*, *Reykðæla saga ok Víga-Skútu*, *Svarfdæla saga*, and *Víglundar saga*. *Vápnfirðinga saga* has occurrences only of *rauðr*, *Þorsteins saga hvíta* and *Þorsteins saga Síðu-Hallssonar* have occurrences only of *hvítr*, *Ljósvefninga saga* has occurrences of *rauðr* and *hvítr*, and *Hænsa-Þóris saga* has occurrences of *rauðr*, *hvítr*, and *svartr*. With toponyms, bynames, all compounds, and derivations of the color adjectives excluded, the representation of the eight basic color terms in the individual Sagas of Icelanders is illustrated in Table 9.

Literary work	rauðr	hvítr	blár	svart	grár	grænn	brúnn	gulr
<i>Bandamanna saga</i>				3			1	
<i>Bárðar saga Snæfellsáss</i>		2		3	6			
<i>Bjarnar saga Hitdælakappa</i>	2	5	4	3				
<i>Droplaugarsona saga</i>				1	1			
<i>Egils saga Skalla-Grímssonar</i>	3	1	5	3	2			
<i>Eiríks saga rauða</i>	4	4	2	3				
<i>Eyrbyggja saga</i>	2	3	4	4	1			1
<i>Finnboga saga ramma</i>	2			1		1		
<i>Fljótsdæla saga</i>	3	3	1	4	3	1	1	
<i>Flóamanna saga</i>	3		3			1		
<i>Fóstbræðra saga</i>	7	9	1	9	1	1		1
<i>Gísla saga Súrssonar</i>		1	3	1	4			
<i>Grettis saga</i>	4	1	4	5	1		3	
<i>Grænlandinga saga</i>		2		1				
<i>Gunnars saga Keldugnúpsfjfls</i>	2			2				
<i>Gunnlaugs saga ormstungu</i>	3	5	1	2	2			
<i>Hallfreðar saga</i>		3			1	1		
<i>Harðar saga ok Hólmverja</i>		1	1	2				
<i>Hávarðar saga Ísfríðings</i>		1	1		1			
<i>Heiðarvíga saga</i>	3	2	1	2	3			
<i>Hrafnkels saga Freysgoða</i>			2					
<i>Hænsa-Þóris saga</i>	1			1				
<i>Kjalnesinga saga</i>	2	1		1				
<i>Kormáks saga</i>	1	1	1	5				
<i>Króka-Refs saga</i>	1	3	3			1		
<i>Laxdæla saga</i>	12	5	9	4	4			1
<i>Ljósvetninga saga</i>	3		1					
<i>Njáls saga</i>	19		9	5	3		3	1
<i>Reykðæla saga ok Víga-Skútu</i>	2	4		1	1	1		
<i>Svarfáæla saga</i>	5		1	1	1			
<i>Þórðar saga hreðu</i>	4	1	1					
<i>Þorskfirðinga saga</i>	2	2						
<i>Þorsteins saga hvíta</i>								
<i>Þorsteins saga Síðu-Halls-sonar</i>								
<i>Valla-Ljóts saga</i>	1		3	1			1	
<i>Vápnfirðinga saga</i>	4							
<i>Vatnsdæla saga</i>	4	2	1	2				
<i>Víga-Glúms saga</i>	1	2	3	1	1	3		
<i>Víglundar saga</i>	1	2					1	

Table 9. Representation of color terms

4 Conclusion

The most striking difference between these sequences and the evolutionary order proposed by Berlin and Kay and modified for Old Norse-Icelandic by Wolf (2006, 2009, and forthcoming) is the prominence of *rauðr*. An explanation for the relative infrequent use of *svartr* and *hvítr* in comparison with *rauðr* may be that in the case of the two achromatic basic color terms there are near-synonyms, such as *dökk* and *myrkr* (for *svartr*) and *ljóss*, *fagr*, and possibly *bleikr* (for *hvítr*), which have not been considered in this study, whereas *rjóðr* is the only near-synonym for *rauðr* and occurs only once (*Kormáks saga* 230.15).⁸

Blár, too, is very common, but, as argued by Wolf (2006: 74), it originally meant simply a dark color and did not attach firmly to the blue spectrum until the late fourteenth century. Considering the fact that the vast majority of the Sagas and *þættir* of Icelanders are believed to have been composed before the late fourteenth century, most of the occurrences of *blár* in these works should, therefore, probably be grouped with *svartr*, which would make *svartr* the most commonly used color term. Accordingly, the order of frequency for the first four (five) color terms is likely as follows: black (*svartr/blár*) > white (*hvítr*) > red (*rauðr*) > grey (*grár*), which is in line with the above proposed sequence for Old Norse-Icelandic.

With regard to the last three color terms, *brúnn*, *grænn*, and *gulr*, the frequency study supports the above proposed sequence in that *gulr* as a color term should be assigned a late stage and after *grænn* and *brúnn*. Indeed, as argued by Wolf (2010: 123), “while yellow (*gulr*) certainly existed, the color was expressed primarily by means of derivatives of *gull* prior to the thirteenth century.”

The dating of the Sagas of Icelanders is a controversial matter, but there seems to be general agreement to date the composition of, for example, *Egils saga Skalla-Grímssonar* to the first half of the thirteenth century, that of *Eyrbyggja saga* and *Laxdæla saga* to around the middle of the thirteenth century, and that of *Njáls saga* and *Hrafnkels saga Freysgoða* to the last quarter of the thirteenth century. Also, most scholars regard *Bárðar saga Snæfellsáss*, *Finnboga saga ramma*, *Fljótsdæla saga*, *Flóamanna saga*, *Grettis saga*, *Harðar saga ok Hólmverja*, *Hávarðar saga Ísfirðings*, *Kjalnesinga saga*, *Króka-Refs saga*, *Svarfdæla saga*, *Þórðar saga hreðu*, and *Víglundar saga* as postclassical sagas.

⁸ This is discounting compounds and derivations of *rjóðr*.

The conclusions are that there seems to be no correlation between the (proposed) date of composition of the individual sagas and their frequency and variety of basic color terms, but that there seems to be a correlation between the frequency of basic color terms in Old Norse-Icelandic and the implicational ordering of those terms according to the Berlin and Kay hierarchy and the amended version for Old Norse-Icelandic.

Literature:

- Berlin, Brent & Eloy Ann Berlin. 1975. Agaruna color categories. *American Ethnologist* 2: 61–87.
- Berlin, Brent & Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
- Íslenszk fornrit. 1933>. Reykjavík: Hið íslenszka fornritafélag.
- Kay, Paul, Brent Berlin & William Merrifield. 1991. Biocultural Implications of Systems of Color Naming. *Journal of Linguistic Anthropology* 1: 12-25.
- Kay, Paul & Chad K. McDaniel. 1978. The Linguistic Significance of the Meanings of Basic Color Terms. *Language* 54: 610-646.
- Regier, Terry, Paul Kay & Naveen Khetarpal. 2007. Color Naming Reflects Optimal Partitions of Color Space. *Proceedings of the National Academy of Sciences of the United States of America* 104: 1436-1441.
- Wolf, Kirsten. 2005. Reflections on the Color of Esau's Pottage of Lentils (*Stjórn* 160.26-161.9). *Gripla* 16: 251-258.
- Wolf, Kirsten. 2006. Some Comments on Old Norse-Icelandic Color terms. *Arkiv för nordisk filologi* 121: 172–192.
- Wolf, Kirsten. 2006. The Color Blue in Old Norse-Icelandic Literature. *Scripta Islandica* 57: 55–78.
- Wolf, Kirsten. 2009. The Color Grey in Old Norse-Icelandic Literature. *Journal of English and Germanic Philology* 108: 222–238.
- Wolf, Kirsten. 2009. Towards a Diachronic Analysis of Old Norse-Icelandic Color Terms: The Cases of Green and Yellow. *Orð og tunga* 11: 109–130.
- Wolf, Kirsten. [Forthcoming]. The Color Brown in Old Norse-Icelandic Literature. In *Altnordische Tradition und christliches Mittelalter: Begegnungen mit den Fremden*. Ed. Kolbrún Haraldsdóttir and Hubert Seelow.

Keywords

basic colors, color terms, color categories, linguistics, quantitative

Lykilorð

grunnlitir, litaorð, litaflokkar, málvísindi, meginleg rannsókn

Útdráttur

Í ritinu *Basic Color Terms* (1969) halda Brent Berlin and Paul Kay því fram að orð um grunnliti komi inn í tungumál í tiltekinni röð: 1: svart, hvítt; 2: svart, hvítt, rautt; 3a: svart, hvítt, rautt, grænt; 3b: svart, hvítt, rautt, gult; 4: svart, hvítt, rautt, gult, grænt; 5: svart, hvítt, rautt, gult, grænt, blátt; 6: svart, hvítt, rautt, gult, grænt, blátt, brúnt; 7: svart, hvítt, rautt, gult, grænt, blátt, brúnt, fjólublátt, bleikt, appelsínugult, grátt. Síðar hafa hugmyndir þeirra verið þróaðar áfram og gagnrýndar nokkuð og til að bregðast við því hafa þessir tveir málvísindamenn og mannfræðingar breytt þeim.

Forníslenska hafði orð um átta grunnliti. Nýlegar rannsóknir á þeim styðja í aðalatriðum þróunarröð Berlin and Kay þótt því hafi verið haldið fram að gera þyrfti ráð fyrir því að *grár* komi snemma fram, að *blár* komi seint fram, að *grænn* tilheyri fyrri stigi en *gulr* og að *brúnn* komi tiltölulega seint fram.

Í þessari grein er kynnt meginleg rannsókn á grunnlitaorðum í Íslendingssögum og Íslendingaþáttum. Markmið hennar var að kanna hvort vensl væru á milli tíðni hugtakanna og þeirrar þróunarraðar sem haldið hefur verið fram, þótt viðurkennt sé að ekki þurfi endilega að vera bein tengsl milli samtímalegrar notkunartíðni og sögulegrar stigskiptingar í tilkomu þeirra. Áberandi munur milli notkunartíðni og þeirrar þróunarraðar sem Berlin og Kay halda fram birtist í því hvað *rauðr* er algengur. Ástæða þess að *svartr* og *hvítr* eru tiltölulega fátíðir í samanburði við *rauðr* kynni að vera sú að þeir fyrnefndu eiga sér ýmis nálæg samheiti en aftur á móti er *rjóðr* eina nálæga samheitið við *rauðr*. *Blár* er líka algengur en þar sem það orð vísaði upprunalega til dökks litar en tengdist ekki bláa litrófinu beint fyrr en seint á 14. öld er væntanlega rétt að flokka stærstan hluta dæmanna um *blár* með *svartr*. Könnun á tíðni *brúnn*, *grænn* og *gulr* styður þá röð sem stungið hefur verið upp á að því leyti að *gulr* tengist síðbúnu stigi og kemur á eftir *grænn* og *brúnn*. Niðurstaðan er sú að jafnvel þótt ekki virðist vera nein vensl á milli (áætlaðs) ritunartíma einstakra sagna og tíðni og fjölbreytileika grunnlitarorða í þeim, virðast vera vensl milli tíðni grunnlitarorða í forníslensku og þróunar þeirra samkvæmt þróunarstigveldi Berlin and Kay í þeirri útgáfu sem löguð hefur verið að forníslensku.

Kirsten Wolf
Department of Scandinavian Studies
University of Wisconsin-Madison
1370 Van Hise Hall
1220 Linden Drive
Madison, Wisconsin 53706
USA
kirstenwolf@wisc.edu



Auður Lorenzo

Íslensk-spænsk orðabók

Guðrún H. Tulinius, Margrét Jónsdóttir, Sigrún Á. Eiríksdóttir, Teodoro Manrique Antón og Viola Miglio (ritstj.), *Íslensk-spænsk orðabók*. Forlagið, Reykjavík. 2011. (xi + 701 bls.) ISBN 978-9979-53-554-6.

1 Inngangur

Haustið 2011 gaf Forlagið út *Íslensk-spænska orðabók (ÍS hér eftir)*. Ritstjórar eru Guðrún H. Tulinius, Margrét Jónsdóttir, Sigrún Á. Eiríksdóttir, Teodoro Manrique Antón og Viola Miglio. Í formálanum kemur fram að við gerð bókarinnar hafi ritstjórar fengið aðstoð annarra og gerð er grein fyrir því í hverju sú aðstoð fólst. Einnig er þess getið að vinna við ÍS hófst árið 2007 og að hún fór öll fram í Lexu, orðabókaforriti Forlagsins, en sama forrit var notað við gerð *Spænsk-íslenskrar orðabókar* sem kom út árið 2007.

Bókin er mikið fagnaðarefni fyrir spænskunemendur á Íslandi sem og alla þá sem eru í einhverjum tengslum við hinn spænskumælandi heim.

Í ÍS eru 27.000 uppflettiorð og ríflega 13.000 orðasambönd og máldæmi. Bókin er rúmlega 700 tvídálka síður, innbundin og þægileg stærð á lettri.

2 Heildarskipan

Fremst í ÍS, framan við flettiorðaskrána, eru formáli, notkunarleið-

Orð og tunga 15 (2013), 163–167. © Stofnun Árna Magnússonar í íslenskum fræðum, Reykjavík.

beiningar og skammstafanaskrá á íslensku. Í formálanum segir frá því hverjir komu að gerð bókarinnar og hvernig hún var fjármögnuð. Einnig er gerð grein fyrir orðaforðanum sem varð fyrir valinu. Í *ÍS* er áhersla lögð á almennan orðaforða úr daglegu nútímamáli en auk þess er þar að finna ýmis sérfræðiorð tengd viðskiptum, hagfræði, jarðfræði, dýra- og náttúrufræði og fleiri greinum. Spænskan á Spáni er lögð til grundvallar en einnig eru í bókinni orð sem eru einkum notuð í Rómönsku Ameríku. Í formálanum kemur einnig fram hver helsti markhópur bókarinnar er. *ÍS* er samkvæmt honum einkum ætluð nemendum á mismunandi skólastigum, sem og áhugamönnum um spænska tungu; það er, bókin virðist hugsuð sem einbeind orðabók handa íslenskum notendum.

Notkunarleiðbeiningar koma næst á eftir formálanum. Sýnd eru dæmi um flettur í bókinni sem sýna hvaða málfræðilegu og merkingalegu upplýsingar koma fyrir í hverri flettu. Skýringar eru mjög mikilvægar þegar uppflettiorðin eiga sér ekki jafnheiti á markmálinu. Þá er orðið útskýrt á markmálinu og er skýringin skáletruð til að auðkenna hana. Almennt er slíkt fyrirkomulag, að útskýra orðið eins og í einmálaorðabókum eða fræðiorðabók, talið mjög mikilvægt í tvímálaorðabókum (Haensch 1982:292). Þar að auki eru gefnar upplýsingar um það ef orð tilheyra sérstöku sérsviði eða málsniði. Slíkt er nauðsynlegt í tvímálaorðabókum því upplýsingar um málsnið gefa vísbendingu um það að viðkomandi orð sé ekki hægt að nota við allar aðstæður. Þá eru orðin t.d. merkt óformlegt, gróft og fleira. Sömuleiðis er í sérstöku tilfellum gerð grein fyrir útbreiðslu eða málsvæðismörkum. Almennt er miðað við spænskuna sem töluð er á Spáni en ef mikill munur er á milli málsvæða er það tekið fram. Þetta er mjög mikilvægt í spænsku vegna þess hversu útbreitt tungumálið er, en spænska er eitt mesta talaða tungumál heims. Sömuleiðis er að finna lista yfir skammstafanir og styttingar sem notaðar eru til þess að tákna ýmis sérsvið. Eins er gerð grein fyrir þeim táknum sem eru notuð í bókinni og hlutverki þeirra og er það gagnlegt fyrir notandann.

Aftur á móti eru engar upplýsingar í formálanum um hvernig orðaforðinn var valinn, þar er einungis greint frá því hvaða forrit var notað til að vinna bókina en ekki hvort stuðst var við tilbúinn orðabókagrunn og ef ekki hvert orðaforðinn var þá sóttur. Einnig er mikilvægt að benda á það að inngangur og notkunarleiðbeiningar eru einungis á íslensku sem verður að teljast galli vegna þess að hér er um að ræða tvímálaorðabók. Einnig ber að nefna að í orðsgreinum er

að finna skammstafarnir sem ekki eru á skammstafanalistanum (þær er hins vegar að finna í *Spænsk-íslenskri orðabók* frá 2007 en þar sem orðabækurnar eru ekki seldar saman dugir það auðvitað ekki).

Í formálanum er einnig gerð grein fyrir markhópnum, eins og áður sagði. Notendahópurinn virðist mjög breiður og erfitt er að uppfylla kröfur allra notenda sem greint er frá í formála í bók eins og *ÍS*. Haensch (1982:397–400) telur að mjög mikilvægt sé að skilgreina vel notendahópin því annars verður eins og í þessu tilfelli erfitt að standast væntingar notenda, því munur er t.d. á þörfum framhaldsskólanemenda og þeirra sem stunda spænskunám á háskólastigi. Í formálanum er einungis talað um notendur sem hafa íslensku að móðurmáli en í notkunarleiðbeiningunum er talað um að ítarlegar málfræðiupplýsingar eigi að hjálpa þeim sem eru að læra íslensku og hafa spænsku sem móðurmál. Því er hægt að segja að hér sé ósamræmi milli formálans og notkunarleiðbeininga og að einhverju leyti sé óljóst hver hinn raunverulegi notandi eigi að vera. Í tvímálaorðabókum ættu upplýsingar um framburð og málfræði að vera á báðum tungumálum ef orðabókin á að geta nýst jafn spænskum sem íslenskum notendum og einnig þyrftu formáli og leiðbeiningar að vera bæði á aðgangsmáli (íslensku) og markmáli (spænsku), sbr. Haensch (1982:514–520). Svo er ekki í *ÍS* og hún er því ófullnægjandi hvað þetta varðar og því erfitt fyrir spænskumælandi notendur að hafa fullt gagn af verkinu.

3 Innskipan

Í flettugreinunum eru gefnar málfræðiupplýsingar um uppflettiorðin, svo sem kyn þeirra og orðflokk, ásamt ýmsum öðrum upplýsingum eins og gerð er grein fyrir í notkunarleiðbeiningum. Í rauninni eru þessar upplýsingar óþarfar fyrir markhópin sem er skilgreindur í formálanum þótt þær séu mikilvægar fyrir spænskumælandi notendur. Í sumun tilfellum eru orðskýringar á íslensku innan sviga sem hjálpa notandanum að greina á milli ólíkra merkinga uppflettiorðsins (Haensch 2004:136). Sömuleiðis eru dæmi um málnotkun þegar við á. Þetta er mjög mikilvægt í tvímálaorðabókum vegna þess hversu gagnlegt það er fyrir notendur að vita hvernig og hvenær á að nota orðið. Í heildina fylgir innskipanin í flestum tilfellum forskrift Haensch um tvímálaorðabækur (1982:514–520). Hins vegar er ýmislegt athugavert við innbyrðis röð upplýsinga. Þegar orð hafa fleiri en eina merkingu er sú sem fyrst er gerð grein fyrir í orðsgreininni ekki alltaf

sú sem er algengust í nútímamáli. Ástæðan gæti verið sú að *Íslensk orðabók* (2002) sé höfð til hliðsjónar en skilgreiningarnar eru oft í sömu röð og í þeirri bók. Í einmálaorðabókum fer röðunin oft eftir því hvaða merking er næst uppruna orðsins (Garriga Escribano 2003:108) en aftur á móti hentar það ekki endilega í tvímálaorðabókum vegna þess að í þeim er allajafna talið eðlilegt að skilgreiningarnar raðist eftir tíðni merkingarinnar, þannig að sú merking sem er algengust sé fyrst í röðinni og þannig koll af kolli. Eins og má sjá á dæmunum hér að neðan er sú ekki raunin í *ÍS*. Í fyrra dæminu vantar þá merkingu sem tengist hárgreiðslu og einnig má velta fyrir sér hvort orðið sé oftast notað um sætabrauð eða snúning. Í seinna dæminu er talið æskilegra að merkingin 'rekkja' komi á undan merkingunni 'svæði, rými'.

snúður snúðs/snúðar, snúðar *kk* • (snúningur) vuelta, giro • (bolla) caracola con glaseado • (RAF) rotor • **dæmi** ▷ *setja á sig snúð* ponerse altivo / *fá e-ð fyrir sinn snúð* ser recompensado por algo

rúm rúms, rúm *hk* • (svæði, rými) espacio ▷ *mér liggur það í léttu rúmi* no me importa mucho, me importa poco / *ryðja sér til rúms* extenderse, generalizarse • (rekkja) cama, lecho ▷ *liggja í rúminu* estar acostado en la cama; (vera veikur) guardar cama

4 Niðurstaða

Loks ber að nefna að stuttu eftir að *ÍS* var gefin út kom í ljós að það vantaði orð sem hefjast á bókstafinum *é* og fyrir vikið vantar, meðal annars, persónuformafnið *ég* í orðabókina.

Hvað sem ýmsum göllum orðabókarinnar líður þá verður ekki annað sagt en að bókin hafi mjög marga kosti og mikilvægt er að vekja athygli á því að *ÍS* sé til og ættu allir sem stunda nám í spænsku að geta nýtt sér þessa bók. Hinsvegar er ráðlagt fyrir næstu útgáfu að endurskoða nokkra þætti í *ÍS* eins og þá sem hér hefur verið bent á.

Heimildir

- Garriga Escribano, C. 2003. La microestructura del diccionario: las informaciones lexicográficas. Í Medina Guerra, A. M. (ritstj.), *Lexicografía española*, bls. 103–126. Barcelona: Editorial Ariel.
- Haensch, G. 1982. *La lexicografía. De la lingüística teórica a la lexicografía práctica*. Madrid: Editorial Gredos.
- Haensch, G. 2004. La lexicografía bilingüe en la actualidad y su valoración. *Revista de Lexicografía, Universidad de La Coruña*, 2003–2004:133–138. http://ruc.udc.es/dspace/bitstream/2183/5482/1/RL_10-7.pdf (12. nóvember 2011).
- Íslensk orðabók*. 2002. (3. útgáfa) Ritstjóri: Mörður Árnarson. Reykjavík: Edda útgáfa.
- Spænsk-íslensk orðabók*. 2007. Ritstjórn: Guðrún H. Tulinius, Margrét Jónsdóttir, Ragnheiður Kristinsdóttir, Sigrún Á. Eiríksdóttir og Teodoro Manrique Antón. Reykjavík: Edda útgáfa.

Auður Lorenzo
Fjölbrotaskólanum í Breiðholti / Háskóla Íslands
aul@fb.is



Bókafregnir

Nýlegar orðabækur

Íslensk samheitaorðabók. Ritstjóri: Svavar Sigmundsson. 3. útgáfa, aukin og bætt. Reykjavík: Styrktarsjóður Þórbergs Þórðarsonar og Margrétar Jónsdóttur/Forlagið. 2012. (xiv + 760 bls.) ISBN 978-9979-53-568-3. (Rafræn útgáfa á *Snara* – vefbókasafni: http://snara.is/bls/um/_samh.aspx.)

Nýlega kom út þriðja útgáfa *Íslenskrar samheitaorðabókar* en hún var fyrst gefin út árið 1985 og aftur þremur árum síðar með smávægilegum endurbótum. Það hefur því orðið talsverð bið eftir endurskoðaðri útgáfu samheitaorðabókarinnar. Svavar Sigmundsson er ritstjóri allra þriggja útgáfanna. Auk prentuðu útgáfunnar er efni samheitaorðabókarinnar nú einnig aðgengilegt í vefbókasafninu Snöru.

Engar grundvallarbreytingar hafa verið gerðar á orðabókinni frá fyrri útgáfum og framsetning flettugreina er með sama sniði og áður. Greinarnar samanstanda af uppflettiorði og einu eða fleiri orðum með sömu eða svipaða merkingu, yfirleitt án frekari skýringa. Málfræðilegar- og merkingarlegar upplýsingar eru í algjöru lágmarki og á það jafnt við um uppfletti-orðin og samheitin. Í einstaka tilvikum eru föst orðasambönd tilgreind í flettugreinunum og samheiti eða samsvaranir við þau og einnig eru stundum gefin andheiti uppflettiorðsins. Dæmi um hvort tveggja má t.d. sjá undir **laust** (samböndin *láta laust*, *sofna laust*, *ekki er laust við*, *á lausu* og andheitið *fast*). Í blaðsíðum talið er nýja útgáfan næstum fjórðungi lengri en sú fyrsta en samkvæmt formála hefur flettum fjölgað úr 44 í 48 þúsund þannig að ætla má að einnig hafi verið aukið við einstakar flettugreinar. Framan við meginhluta orðabókarinnar eru birtir formálar allra útgáfanna ásamt skrá um helstu skammstafanir en aftur á móti eru engar leiðbeiningar um uppbyggingu og notkun orðabókarinnar fyrir utan þær upplýsingar sem gefnar eru í einhverjum formálanna þriggja.

Íslensk-ensk, ensk-íslensk vasaorðabók. Ritstjórar: Laufey Leifsdóttir og Una Guðlaug Sveinsdóttir. Reykjavík: Forlagið. 2012. (1037 bls.) ISBN 978-9979-53-567-6.

Þessi orðabók er byggð á tveimur nýlegum orðabókum milli íslensku og ensku, *Íslensk-enskri orðabók* (2009; ritstj. Christopher Sanders o.fl.) og *Ensk-íslensku orðabókinni* (2006; ritstj. Jón Skaptason o.fl.), og má segja að hún sé smækkuð útgáfa af þeim í tvennum skilningi. Eins og nafnið bendir til er hún í litlu broti (þótt hún fari kannski ekki rétt vel í vasa vegna þykktarinnar) og efniviðurinn er minni en í hinum bókunum, um 40.000 uppflettiorðsamkvæmt upplýsingum á baksíðu en samanlagt mun flettufjöldinn í hinum bókunum tveimur vera nálægt 64.000. Ensk-íslenski hlutinn er talsvert lengri og efnis-meiri en sá íslensk-enski, m.a. vegna þess að þar eru ítarlegri upplýsingar um málfræðileg atriði „enda gert ráð fyrir að það henti skólafólki best“ eins og segir í formála. Sá hópur virðist helsti markhópur bókarinnar (þótt á baksíðu komi fram að hún sé einnig hentug fyrir almenning) og ljóst er að þar er um íslenska skólanemendur að ræða því allt stoðefni er eingöngu á íslensku (formáli og notkunarleiddbeiningar ásamt skammstöfunum og skýringum í orðsgreinum í báðum hlutum orðabókarinnar). Lauslegur samanburður bendir til þess að styttingar í báðum hlutum felist annars vegar í því að sleppa flettugreinum (gjarnan samsettum og afleiddum orðum sem telja má tiltölulega gagnsæ og/eða óalgeng, t.d. *atkvæðaskipting*, *kauphallarviðskipti*; *local colour*, *localization*) og hins vegar í því að stytta greinar, t.d. með því að fella niður dæmi. Auk þess eru engar framburðarupplýsingar í bókinni öfugt við *Ensk-íslensku orðabókina* þar sem öll uppflettiorð eru hljóðrituð. Stoðefni er líka stillt í hóf – formálinn er stuttur, notkunarleiddbeiningar hnitmiðaðar og skrá um tákn og skammstafanir eru á einni síðu á eftir þeim; sú skrá er þó jafnframt birt innan á kápu (baksíðu) þannig að auðvelt er að finna þær. Aftan við íslensk-enska hlutann er listi yfir sterkar og óreglulegar enskar sagnir og beygingu þeirra.

SignWiki Ísland. Tákn málsorðabók og þekkingarbrunnur. Reykjavík: Samskiptamiðstöð heyrnalausra og heyrnarskertra (SHH): <http://is.signwiki.org/index.php/Forsíða>. (3.964 tákn; 25. febrúar 2013)

Á vegum Samskiptamiðstöðvar heyrnalausra og heyrnarskertra (SHH) hefur verið rekin vefsíða frá 2011 þar sem aðgangur er að íslenskri tákn málsorðabók auk kennsluefnis, æfinga og annars fróðleiks um íslenska tákn málið. Allt efnið er aðgengilegt í snjallsímum og spjaldtölvum auk hefðbundinna tölva og er því tiltækt þeim sem á þurfa að halda í dagsins önn. Táknnum er jafnt og þétt bætt við orðabókina og þegar þetta er skrifað eru þau tæplega fjögur þúsund. Notendur geta sjálfir lagt til efni og tákn og þeir geta jafnframt gert tillögur um breytingar á því sem þegar er á síðunni.

Hægt er að nálgast táknin í orðabókinni frá nokkrum hliðum. Aðal-

aðgangurinn er eftir íslenskum orðum eða hugtökum sem birtast í heild í stafrófsröð í flokknum „Orðabók A-Ö“ en einnig er hægt að leita að tilteknu hugtaki. Jafnframt má leita að táknum eftir málfræðilegum einkennum þeirra í flokknum „Málfræðileit“; þar byggjast fyrirspurnirnar á þremur sviðum: handformi, myndunarstað og orðflokki og má velja málfræðileg einkenni af einu, tveimur eða öllum þremur sviðunum. Loks er hægt að nálgast táknið eftir undirflokkum, bæði merkingarsviðum (t.d. „Athafnir“ eða „Eldhús og eldamennska“) og málfræðilegum einkennum (t.d. einstökum orðflokkum eða hlutverki, s.s. spurningum).

Hvert tákn í orðabókinni hefur sína síðu þar sem annars vegar er myndband með tákni og hins vegar rammi með ýmiss konar upplýsingum um táknið. Í honum er kyrramynd af tákni og teikningar sem sýna flokkun þess eftir myndunarstað og handformi. Þar fyrir neðan eru málfræðilegar upplýsingar um táknið í orðum (orðflokkur, myndunarstaður og handform) og auk þess efnisflokkur, upplýsingar um munnhreyfingar og vísun til skyldra tákna sem finna má í orðabókinni. Alls staðar eru tenglar þannig að hægt er að nálgast önnur tákn af sama flokki, t.d. öll tákn með sama handform eða í sama efnisflokki.

Rit um orðabókafræði

Eaker, Birgit, Lennart Larsson og Anki Mattisson (ritstj.). *Nordiska studier i lexikografi 11*. Rapport från Konferensen om lexikografi i Norden, Lund 24–27 maj 2011. (Skrifter utgivna af Nordiska föreningen för lexikografi. Skrift nr 12.) Lund: NFL, Språkrådet i Norge og Svenska Akademiens ordboksredaktion. 2012. (628 bls.) ISBN 978-91-85333-42-4. ISSN 083-9313.

Bókin er ráðstefnurit frá elleftu norrænu orðabókafræðiráðstefnunni sem haldin var í Lundi vorið 2011. Þar eru birtar rúmlega fimmtíu greinar sem byggðar eru á fyrirlestrum sem haldnir voru á ráðstefnunni, þ. á m. fyrirlestrum þriggja boðsfyrirlesara – Godelieve Laureys frá Gent í Belgíu („Mot en hållbar lexikografi – En balansgång mellan återanvändning och förnyelse“), Bo Ralph frá Gautaborg („Biskopar, ordböcker och verkligheten – Vad är lexikografihistoria bra för?“) og John Simpson frá Oxford, forstöðumanns við Oxford English Dictionary („What has the OED become?“). Aðrar greinar, sem flestar eru skrifaðar af norrænum orðabókafræðingum, birta gott yfirlit yfir orðabókaverkefni sem unnið er að á Norðurlöndum og ýmis viðfangsefni sem fólk þarf að kljást við í því sambandi. Allmargir íslenskir fræðimenn eiga greinar í ritinu og endurspegla viðfangsefni þeirra fjölbreytnina í bókinni almennt: Kristín Bjarnadóttir fjallar um beygingargagnagrunn og tengingu hans við rafræna orðabók; Helgi Haraldsson skrifar um íslensk-tékkneska/tékknesk-íslenska orðabók sem hann er með í

smíðum ásamt Mörtu Bartošková í Prag; Gunnlaugur Ingólfsson ritar grein um fyrstu prentuðu íslensku orðabækurnar; Jón Hilmar Jónsson fjallar um atviksorð og atviksliði og meðhöndlun þeirra í orðabókum; Guðrún Kvaran skrifar um nokkur óprentuð íslensk orðabókahandrit frá 18. og 19. öld sem heimildir fyrir sögulegar orðabækur; loks fjalla Halldóra Jónsdóttir og Þórdís Úlfarsdóttir um hina nýju ISLEX-veforðabók sem hefur íslensku sem viðfangsmál og nokkur norræn markmál. ISLEX-orðabókin kemur einnig við sögu í grein Ylva Hellerud sem fjallar um tvímálaorðabækur frá sjónarhóli þýðanda og í grein þeirra Margunn Rauset, Önnu Helgu Hannesdóttur og Aldísar Sigurðardóttur („Ein-, to- eller fleirspråkleg ordbok?“), en allir þessir höfundar hafa komið að ritstjórnarvinnu í verkinu.

Fjeld, Ruth Vatvedt, og Julie Matilde Torjusen (ritstj.). *Proceedings of the 15th EURALEX International Congress*. Oslo: Department of Linguistics and Scandinavian Studies, UiO. (xii + 244 bls.) ISBN 978-82-303-2095-2. (Einnig á vefnum: http://www.euralex.org/proceedings-toc/euralex_2012/).

Fimmtánda ráðstefna EURALEX, evrópskra samtaka orðabókafræðinga, var haldin í Osló sumarið 2012 og samhliða var gefið út ráðstefnurit. Það er tvískipt. Annars vegar er prentað rit þar sem birtar eru greinar boðsfyrirlesaranna ásamt útdráttum úr öðrum fyrirlestrum, kynningum á hugbúnaði og veggspjöldum. Hins vegar er rafræn útgáfa sem bæði fylgir ritinu á minnislykli og er birt á vefsíðu EURALEX-samtakanna en þar má nálgast greinar úr öllum fyrirlestrum sem haldnir voru á þinginu.

Meginþema ráðstefnunnar var *Orðabókafræði og þjóðarímynd* („Lexicography and National Identity“) en auk þess var sjónum sérstaklega beint að eftirtöldum sex sviðum: (1) Frumbyggjamál og orðabókafræði; (2) Málheildir og orðabókafræði; (3) Orðabókafræði í máltækni; (4) Margmála orðabókafræði; (5) Orðabókafræði og kenningar í merkingarfræði; (6) Íðorðfræði, sérfræðimál og orðabókafræði; (7) Skýrslur um orðabókarverk og orðfræðileg verkefni. Öll þessi efni endurspeglast í útdráttum og greinum sem birt eru í ráðstefnuritinu. Fimm boðsfyrirlestrar voru fluttir á þinginu og efni þeirra tengist ofantöldum þemum. Eins og áður segir birtast greinar unnar upp úr þeim í prentaðri útgáfu ráðstefnuritsins. Framlag Ole Henrik Magga nefnist „Lexicography and indigenous languages“ og fjallar um samísku og samískar orðabækur. Arnfinn Muruvik Vonen ræðir norsk ritmálsafræði í greininni „Diversity and democracy: written varieties of Norwegian“. Grein Bolette Sandford Pedersen hefur titilinn „Lexicography in Language Technology (LT)“ og snýst um tengsl orðabókafræði og máltækni. Michael Rundell fjallar um orðabókagerð og kenningar í orðabókafræði undir yfirskriftinni „It works in practice but will it work in theory? The uneasy relationship between lexicography and matters theoretical“. Loks gerir Gilles-Maurice de Schryver grein fyrir straumum og stefnum í orðabókafræði í grein sem

nefnist „Lexicography in the Crystal Ball: Facts, trends and outlook“ þar sem hann styðst við greiningu á umfjöllunarefnum á fyrri ráðstefnum EURALEX. Útdrættir úr erindum eru allir prentaðir í bókinni og þar er þeim raðað eftir efnisflokkum sem svara n.v. til þeirra þema sem talin voru að ofan. Aftan við þá er efnisorðaskrá og skrá yfir alla höfunda. Eins og áður sagði verður að nálgast greinar úr flestum fyrirlestrunum í rafrænni útgáfu ráðstefnuritsins eða á vefsíðu samtakanna.

Granger, Sylviane, og Magali Paquot (ritstj.). *Electronic Lexicography*. Oxford: Oxford University Press. 2012. (xiii + 517 bls.) ISBN 978-0-19-965486-4.

Hér er á ferðinni nýtt greinasafn um stöðu rannsókna og þróunar í rafrænni orðabókafræði, þ.e.a.s. notkun tölvutækni við skipulagningu, gerð, dreifingu og notkun orðabóka. Ritstjórnirnar starfa báðir við háskólann í Louvain í Belgíu og höfundar greinanna níttán sem birtar eru í safninu eru starfandi orðabókafræðingar við háskóla, rannsóknarstofnanir og fyrirtæki, aðallega í Evrópu en sumir koma frá Suður-Afríku og Kanada.

Annar ritstjórnanna, Sylviane Granger, skrifar ítarlegan inngang að safninu þar sem hún fjallar um rafræna orðabókafræði eins og hún birtist í ritinu og þær róttæku breytingar sem hafa orðið í greininni frá því að tölvutækni var fyrst innleidd við orðabókagerð. Þar telur hún sex nýjungar skipta mestu máli fyrir þróun orðabókafræði og úrlausnarefni á því sviði í nánustu framtíð: (1) tilkomu málheilda og notkun þeirra; (2) meiri og betri heimildir en áður; (3) mikilvægi aðgengis að efni orðabóka; (4) aðlögun að ólíkum þörfum; (5) samruna orðabóka og annarra málgagna; (6) aðkomu notenda að orðabókagerð. Auk þess gerir hún grein fyrir ritinu og greinunum sem þar birtast og loks dregur hún saman helstu þræði og gefur stutt yfirlit yfir mikilvægustu verkefni á sviði rafrænnar orðabókafræði á næstu árum.

Greinunum er skipað í þrjá nokkurn veginn jafnstóra hluta. Í þeim fyrsta, „Lexicography at a watershed“, eru greinar sem fjalla um gögn og töl til orðabókagerðar og áhrif þeirra á hugmyndir og kenningar í orðabókafræði. Sá næsti hefur yfirskriftina „Innovative dictionary projects“ og þar er gerð grein fyrir orðabókaverkefnum sem fela í sér ýmiss konar nýjungar, m.a. með samruna við annars konar málöggn og heimildir. Í þriðja hlutanum, „Electronic dictionaries and their users“, er áherslan síðan á orðabókanotendur og orðabókanoftkun við ólíkar aðstæður. Í lok bókarinnar eru svo gagnlegar skrár um heimildir, höfunda og atriðisorð.

Meðferð íslensku fyrr og síðar

Stefán Pálsson, Anton Kaldal Ágústsson, Gunnar Vilhjálmsson og Steinar Ingi Farestveit. *ð ævisaga*. Reykjavík: Crymogeia. 2012. (199 bls.) ISBN 978-9935-420-26-8.

Í bókinni er rakinn uppruni og saga bókstafsins *ð* sem er einn af hinum sér-íslensku stöfum í stafrófinu þótt ekki sé hann heimaþinginn í upphafi heldur kominn úr engilsaxneskri rithefð á árdögum ritaldar á Íslandi. Sjórnarhornið er einkum á notkun hans í íslensku ritmáli fyrr og síðar, hvarfi hans í upphafi prentaldar og endurupptöku á 19. öld. Það var ekki síst Rasmus Kristian Rask sem stuðlaði að endurvakningu *ð*-sins á þeim tíma og ekki voru allir samtímamenn hans jafnsannfærðir um nauðsyn bókstafsins og ágæti. En ekki er einungis staldrað við sögu og málsögu heldur er einnig fjallað um bókstafinn *ð* frá leturfræðilegu sjónarmiði, gerð grein fyrir grafískum einkennum hans og glímunni við að fella hann að latneskum leturgerðum fyrr og síðar og loks baráttunni um að fá hann tekinn upp í alþjóðlega leturstaðla. Höfundar bókarinnar koma úr ólíkum áttum, einn þeirra er sagnfræðingur og hinir þrír grafískir hönnuðir, og stuðlar það eflaust að því hvað farið er víða í umfjölluninni. Bókin er ríkulega myndskreytt og allur ytri umbúnaður hennar ber þess merki að þar áttu hönnuðir hlut að máli. Það vekur líka athygli að hún er gefin út af forlagi sem er ekki síst þekkt fyrir útgáfu myndlistar- og ljósmyndabóka enda er hinn sjónræni þáttur ekki síður mikilvægur þáttur í ritinu en textinn.

Eiríkur Rögnvaldsson, Kristín M. Jóhannsdóttir, Sigrún Helgadóttir og Steinþór Steingrímsson. *Íslensk tunga á stafrænni öld / The Icelandic Language in the Digital Age*. (Hvítbókaröð METANET.) Berlin/Heidelberg: Springer Verlag. 2012. (73 bls.) ISBN 978-36-42301-73-5. ISSN 2194-1416.

Rit þetta er svonefnd *hvítbók* sem gefin er út í tengslum við verkefnið METANET, sem Íslendingar og aðrir Norður-Evrópubúar eiga aðild að undir merkjum METANORD, og útgáfa hennar er kostuð af áætlunum Evrópusambandsins í upplýsinga- og samskiptatækni. Hafa sambærileg rit verið gefin út fyrir önnur tungumál í álfunni. Tilgangurinn með þessum ritum er að lýsa stöðu málanna og möguleikum gagnvart nýjungum í upplýsingatækni og gera grein fyrir því sem unnið hefur verið í máltækni fyrir viðkomandi tungumál. Þar er stuðst við samræmd viðmið til þess að meta ýmsa þætti í stöðu mála á því sviði. Ritið skiptist í fimm meginkafla. Að loknum yfirlitskafla er fyrst fjallað um hættur sem steðja að tungumálinu í stafrænum heimi, í næsta kafla er stutt lýsing á helstu einkennum íslensks máls og málsamfélags, þar á eftir er kafla sem fjallar um máltækni fyrir íslensku þar sem gerð er grein fyrir því sem þegar hefur verið unnið á því sviði og loks er kafla sem segir frá METANET. Til þess að kanna

stöðu máltækni á Íslandi voru fjórir þættir athugaðir sérstaklega: Talvinnsla, vélþýðingar, textagreining og málföng. Niðurstaðan var sú að þrátt fyrir það sem áunnist hefur undanfarinn áratug standi íslenska höllum fæti á öllum þessum sviðum í samanburði við önnur Evrópumál því hún fellur alls staðar í flokkinn „Lítill/enginn stuðningur“ ásamt (mismörgum) öðrum málum sem tiltölulega fáir tala. Bókin er tvímála því textinn er bæði á íslensku og ensku og henni er því bæði beint inn á við – til stjórnvalda og almennings á Íslandi – og út á við, ekki síst til samstarfsaðila innan verkefnisins.

Ný bók um færeyskt málsamfélag

Jógvan í Lon Jacobsen. *Ærligt talt, who cares? En sociolingvistisk undersøgelse af holdninger til og brug af importord og afløsningsord i færøsk*. (Moderne importord i språka i Norden XIII.) Oslo: Novus forlag. 2012. (324 bls.) ISBN 978-82-7099-682-7.

Bók Jógvans í Lon Jacobsen er þrettánda ritið í ritröðinni „Moderne importord i språka i Norden“ þar sem gerð er grein fyrir niðurstöðum stórrar sam-
anburðarrannsóknar á aðkomuorðum í Norðurlandamálum, notkun þeirra og viðhorfum málnotenda til erlendra áhrifa. Einkum var unnið að því verkefni á árunum 2000–2005. (Veturliði Óskarsson ritaði umsögn um ritröðina í *Orð og tungu* 14/2012.) Ritið er dönsk þýðing á doktorsritgerð Jógvans frá 2008 með smávægilegum endurbótum. Í henni rannsakaði hann viðhorf færeyskra málnotenda til erlendra áhrifa, einkum aðkomuorða, í færeysku og beitti til þess ólíkum aðferðum (skoðanakönnun, viðtölum og svonefndu grímuprófi) eins og gerð er grein fyrir í ritinu. Í ritinu er fjallað ítarlega um niðurstöður færeysku rannsóknanna og þær eru jafnframt bornar saman við niðurstöður sambærilegra rannsókna í öðrum norrænum málsamfélögum, þ. á m. á Íslandi. Niðurstöðukaflinn er bæði birtur á dönsku og ensku.

Mál og málvísindi

Herbst, Thomas, Susen Faulhaber og Peter Uhrig (ritstj.). *The Phraseological View of Language. A Tribute to John Sinclair*. Berlin/Boston: De Gruyter Mouton. 2011. (xii + 324 bls.) ISBN 978-3-11-025688-8.

Þetta rit er greinasafn sem gefið var út til heiðurs og í minningu breska málvísindamannsins John McHardy Sinclair (1933–2007) og á rætur að rekja til málþings sem haldið var í tilefni þess að hann var gerður að heiðursdoktor við háskólann í Erlangen í Þýskalandi 2007. Yfirskrift málþingsins var *Chunks in Corpus Linguistics and Cognitive Linguistics*. Sinclair lést áður en til málþingsins kom og þannig varð bókin, sem geymir greinar sem byggjast á

erindum sem þar voru haldin, að minningarriti. Efni greinanna ber svip af áhugasviðum Sinclair innan málvísindanna. Hann var prófessor í ensku við háskólann í Birmingham og er líklega þekktastur fyrir að vera helsti hugmyndasmiður og aðalritstjóri COBUILD-orðabókarinnar sem kom út 1987 og markaði tímamót í orðabókagerð á fleiri en einn veg. Þetta var t.d. fyrsta orðabókin sem var samin algjörlega á grundvelli afmarkaðrar málheildar (COBUILD = *The Collins Birmingham University International Language Database*) en Sinclair stjórnaði gerð hennar og var þannig einn af frumkvöðlum gagnamálfræðinnar (e. *corpus linguistics*). Þriðja áhugasvið Sinclair sem setur svip sinn á ritið er orðræðugreining (e. *discourse analysis*) og málnotkun.

Greinunum í ritinu er skipað í fjóra flokka. Í fyrsta hluta eru, auk formála, tvær greinar sem fjalla um viðfangsefni og feril Sinclair, önnur eftir Michael Stubbs og hin eftir Stig Johansson (sem sjálfur féll frá áður en bókin kom út). Í öðrum hluta eru sex greinar sem tengjast hugtakinu orðastæða (e. *collocation*) á ýmsan hátt, bæði frá hagnýtu og fræðilegu sjónarhorni. Þriðji hlutinn geymir fjórar greinar um tilbrigði og málbreytingar og í fjórða og síðasta hlutanum eru þrjár greinar sem tengjast gagnamálfræði og notkun málheilda, einkum við greiningu á föstum orðasamböndum.

Andersen, Gisle (ritstj.). *Exploring Newspaper Language. Using the web to create and investigate a large corpus of modern Norwegian*. (Studies in Corpus Linguistics 49.) Amsterdam/Philadelphia: John Benjamins Publishing Company. 2012. (356 bls.) ISBN 978-90-272-0354-0.

Þetta greinasafn fjallar um gerð málheildar með norskum dagblaðatextum á vefnum. Í ítarlegum inngangskafla þeirra Gisle Andersen og Knut Hofland er fjallað um byggingu norsku dagblaðamálheildarinnar sem hófst 1998 (*Norsk aviskorpus*; sjá <http://avis.uib.no/>). Textarnir eru sóttir sjálfvirkt frá degi til dags í vefútgáfur norskra blaða. Í innganginum er jafnframt gerð grein fyrir ritinu og tilgangi þess sem er bæði að lýsa aðferðum og tólum, sem þróuð hafa verið í tengslum við verkefnið, og að kynna niðurstöður rannsókna sem byggja á málheildinni. Í fyrri hluta bókarinnar eru sex greinar um aðferðir og tól til þess að nýta vefinn til þess að setja saman málheild. Þar er fjallað um málfræðilega og setningafræðilega mörkun textanna og tól sem nýta má til slíkrar greiningar, um orðasambönd í textunum, um sjálfvirka flokkun eftir umræðuefnum og um aðferðir til þess að bera kennsl á ensk aðkomuorð í textunum. Í síðari hlutanum eru átta greinar þar sem sagt er frá málrannsóknum sem byggjast á málheildinni, þ. á m. um athugun á aðlögun aðkomuorða, um nýyrði í nútímanorsku, um samsetningar með liðnum *aske* á ákveðnu tímabili í kjölfar gossins í Eyjafjallajökli og um myndhverfingar sem tengjast rými og áttum. Höfundar greinanna eru flestir vel þekktir norskir málvísindamenn, einkum á sviði gagnamálfræði og máltækni en einnig orðabókafræði.

Kortmann, Bernd, og Johan van der Auwera (ritstj.). *The Languages and Linguistics of Europe. A Comprehensive Guide*. (The World of Linguistics 1.) Berlin/Boston: De Gruyter Mouton. 2011. (xviii + 911 bls.) ISBN 978-3-11-022025-4.

Þetta mikla greinasafn um mál og málvísindi í Evrópu er fyrsta ritið í ritröð um heim málvísindanna (áætlað er að næstu tvö rit, um frumbyggjamál í Suður-Ameríku og mál og málvísindi í Ástralíu, komi út 2013). Ritstjórnarnir gera grein fyrir ritinu í formála en síðan er það samsett af fjölda greina eftir evrópska málvísindamenn. Ritinu er skipt í fimm meginhluta. Í þeim fyrsta eru fjórtán greinar um málgerð (e. *typology*) evrópskra tungumála sem töluð eru allt austur að Úralfjöllum og Kákasus. Þær fjalla ýmist um málfræðileg einkenni tiltekinna málaætta (t.d. keltneskra og germanskra mála) eða einstakra mála (t.d. albönsku og basknesku) frá samtímalegu sjónarhorni. Síðasta greinin í þessum hluta fjallar svo um evrópsk táknmál. Næsti hluti fjallar um svæðisbundna málgerð og máltengsl og skiptist hann í fjóra undirflokkka: „Areal typology“ með fjórum greinum um tiltekin málsvæði, þ. á m. Balkanskaga og Miðjarðarhafssvæðið; „Language contact“ með fimm greinum um máltengsl á ákveðnum svæðum, þ. á m. grein eftir finnska málvísindamanninn Jan-Ola Östmann um Norður-Evrópu, einkum Norðurlönd; „Minority languages“ með tveimur greinum – annars vegar um gömul minnihlutamál í Evrópu og hins vegar um innflytjendamál í álfunni; og loks „Non-standard varieties“ með þremur greinum sem fjalla á einn eða annan hátt um mállýskur, staðalmál og mállýskutengsl. Í þriðja hluta eru greinar um málpólítík og málstefnu í Evrópu. Þar eru átta greinar sem fjalla um slík efni frá ýmsum hliðum, t.d. út frá máltengslum í tví- eða fjöltyngdum málsamfélögum, frá femínískum sjónarhóli og með hliðsjón af enskum áhrifum á önnur Evrópumál. Fjórði hluti bókarinnar geymir sömuleiðis átta greinar og þær fjalla um sögu evrópskra tungumála. Þar er m.a. fjallað um áhrif fólksflutninga á málalandslag Evrópu, fjöltyngd heimsveldi og áhrif þeirra á sögu álfunnar (Habsborgararíkið, Ottómanaveldið, Rússneska keisaradæmið og Sovétríkin) og samband tungumáls og myndunar þjóðríkja. Í fimmta og síðasta hlutanum er svo sjónum beint að ástundun og þróun málvísinda og málrannsókna í Evrópu. Þar eru fimm greinar. Sú fyrsta fjallar um rannsóknarhefðir fram til 18. aldar og sú næsta um málrannsóknir og kenningar í málvísindum á 19. öld. Þrjár síðustu greinarnar fjalla um þrjár útbreiddar stefnur í málvísindum á 20. öld og ástundun þeirra í Evrópu: strúktúralisma eða formgerðarstefnu, fúnksjónalisma og generatíf málvísindi. Bókinni lýkur með ítarlegum atriðisorðaskráum yfir tungumál og málbrigði, nöfn og efnisatriði.



Ráðstefnur 2013

Fjölmargar áhugaverðar norrænar og alþjóðlegar ráðstefnur um mál og málfræði, orðabókafræði, nafnfræði og málrækt verða haldnar á árinu. Hér á eftir eru nokkrar þær helstu taldar upp í tímaröð og gerð stuttlega grein fyrir viðfangsefni hverrar um sig.

Norræn málvísindaráðstefna í Reykjavík

25. norræna málvísindaráðstefnan (*25th Scandinavian Conference of Linguistics*) verður haldin í Háskóla Íslands dagana 11.–13. maí undir merkjum samtaka norrænna málvísindamanna (NAL: *Nordic Association of Linguists*). Gestafyrirlesarar verða Elisabeth Engberg-Pedersen, Hans-Martin Gärtner og Alec Marantz. Tólf málstofur hafa verið auglýstar en auk þess verður fluttur fjöldi annarra fyrirlestra um margvísleg málvísindaleg efni. Ítarlegar upplýsingar um ráðstefnuna, m.a. um skráningu á hana, er á vefsíðunni <http://conference.hi.is/scl25/>.

Máltækniráðstefna í Osló

NODALIDA 2013, 19. norræna ráðstefnan um tölvumálvísindi (e. *computational linguistics*), verður haldin á vegum *The Northern European Association for Language Technology* (NEALT) í Háskólanum í Osló 22.–24. maí. Þrjú boðsfyrirlestrar verða fluttir á ráðstefnunni, fimm málstofur hafa verið skipulagðar og fjöldi annarra fyrirlestra á sviði tölvumálvísinda og máltækni verða í boði. Allar upplýsingar um ráðstefnuna má nálgast á vefsíðu hennar: <http://www.hf.uio.no/iln/english/research/news-and-events/events/conferences/2013/nodalida/index.html>.

Ráðstefna um málstöðlun og forskrift

Ráðstefna um forskrift og fyrirmæli um mál og málnotkun – *Prescriptivism Conference 2013* – verður haldin í háskólanum í Leiden í Hollandi 12.–14. júní. Þetta er í fjórða sinn sem slík ráðstefna er haldin og í þetta sinn er þemað „Forskrift og hefð í máli“ (*Prescription and Tradition in Language*). Frekari upplýsingar eru á vefsíðu ráðstefnunnar: <http://hum.leiden.edu/lucl/prescriptivism-conference/>

Ráðstefna um máltilbrigði í Evrópu

Sjöunda alþjóðlega ráðstefnan um máltilbrigði í Evrópu (ICLaVE: *International Conference on Language Variation in Europe*) verður haldin í háskólanum í Þrándheimi 26.–28. júní. Auglýstar hafa verið fimm málstofur auk stakra fyrirlestra og þriggja boðsfyrirlestra. Á ráðstefnunni verður fjallað um ýmis efni sem tengjast rannsóknum á máltilbrigðum og meðal þátttakenda verða málvísindamenn á sviði sögulegrar málfræði, félagsmálfræði og mállýskufraði. Frekari upplýsingar eru á vefsíðu ráðstefnunnar: <http://hist.no/iclave7/>.

Ráðstefna um söguleg málvísindi

21. alþjóðlega ráðstefnan um söguleg málvísindi (ICHL: *International Conference on Historical Linguistics*) verður haldin í Háskólanum í Osló 5.–9. ágúst. Fimm gestafyrirlestrar verða fluttir á ráðstefnunni auk fjölda annarra fyrirlestra og auglýstar hafa verið sex málstofur um fjölbreytileg efni. Ítarlegar upplýsingar eru um ráðstefnuna á vefsíðunni <http://www.hf.uio.no/ifikk/english/research/events/ichl2013/>.

Norræn ráðstefna um orðabókafræði

Tólfta ráðstefnan um orðabókafræði á Norðurlöndum verður haldin í nafni NFL (*Nordisk forening for leksikografi*) í Osló 13.–16. ágúst. Þrjú gestafyrirlesarar halda erindi á ráðstefnunni: Arleta Adamska, Victoria Rosén og Lars Trap-Jensen. Aðrir fyrirlestrar munu fjalla um ólík efni á sviði orðabókafræði, fræðilegar jafnt sem hagnýtrar. Frekari upplýsingar má nálgast á vefsíðunni <http://nordisksprogkoordination.org/nfl/konferencer-og-symposier/12.-konferanse-om-leksikografi-i-norden>.

Frændafundur

Færeysk-íslenska málþingið *Frændafundur* verður haldið í áttunda skipti 24.–25. ágúst og í þetta sinn fer það fram í Þórshöfn. Málþingið er þverfaglegt og óskað er eftir fyrirlestrum af ólíkum fræðasviðum um efni sem eru áhugaverð fyrir Færeyjar og Ísland. Frekari upplýsingar eru á vefsíðunni <http://setur.fo/foeroyamalsdeildin/fraendafundur-2013/>.

Þing norrænu málnefndanna

Norrænt málnefndaþing (*Nordisk sprogmode*) verður haldið á Akureyri 28.–29. ágúst 2013. Þema þingsins er *De nordiske sprogs fremtid indenfor informa-*

tionsteknologien. Máltækni verður sífellt mikilvægari en tungumál sem fáir tala, eins og norrænu málin, standa höllum fæti gagnvart henni. Á þinginu verður rætt um hver staða tungumála á Norðurlöndum er gagnvart nýrri tækni og hvernig bæta megi stöðuna. Þing norrænu málnefndanna er haldið árlega og var síðast haldið á Íslandi árið 2006.

Norrænt málþing um nafnfræði

Málþing NORNA, norrænna samtaka um nafnfræðirannsóknir, verður haldið í 43. skipti í Halmstad í Svíþjóð 6.–8. nóvember. Þema málþingsins að þessu sinni er „Nýjungar í nöfnum og nafnavenjum“ (*Innovationer i namn och namn-mönster*). Frekari upplýsingar um þingið má finna á vef samtakanna: <http://www2.sofi.se/NORNA/NORNA.html>.